



UNIVERSITÀ DEGLI STUDI DI PALERMO

Dottorato di Ricerca in Matematica e Informatica

Dipartimento di Matematica e Informatica

Settore Scientifico Disciplinare INF/01

METODOLOGIE E TECNICHE PER L'ANALISI DI DATI COMPLESSI IN AMBIENTE ECOGRAFICO PRENATALE

IL DOTTORE
GIUSEPPA SCIORTINO

IL COORDINATORE
LUISA DI PIAZZA

IL TUTOR
DOMENICO TEGOLO

CICLO XXVI
ANNO CONSEGUIMENTO TITOLO 2016

Indice

Indice	1
Introduzione	3
1 Ultrasuoni ed ecografia prenatale	6
1.1 Gli ultrasuoni	7
1.1.1 I sistemi ecografici	10
1.1.2 Sistemi di acquisizione	11
1.1.3 Sonde, risoluzioni spaziali	11
1.1.4 Modalità di rappresentazione	12
1.2 Medical imaging mediante ultrasuoni	15
1.2.1 Ecografia	15
1.2.2 Artefatti	17
1.2.3 Ecografia fetale	21
1.2.4 La traslucenza nucale e le anomalie cromosomiche ad essa associata	25
1.2.5 La misurazione della traslucenza nucale	28
1.2.6 Sezione sagittale mediana e traslucenza nucale	30
2 Metodi e tecniche	34
2.1 Trasformata wavelet	35
2.1.1 Trasformata di Fourier	36
2.1.2 Short Time Fourier Transform	37
2.1.3 La trasformata wavelet continua	39
2.1.4 Trasformata wavelet discreta	42

2.1.5	Trasformata wavelet veloce	47
2.1.6	Caso bidimensionale	50
2.1.7	La trasformata à <i>trous</i>	53
2.2	Reti Neurali Artificiali	59
2.2.1	Reti multistrato	66
2.2.2	Reti neurali e problemi di classificazione	74
2.3	Trasformata simmetria	76
3	Metodologia	80
3.1	Stato dell'arte	80
3.2	Applicativo	86
3.2.1	Individuazione della sezione sagittale mediana	87
3.2.2	Individuazione della translucenza nucale	98
3.2.3	Misurazione della translucenza nucale	100
4	Risultati	103
4.1	Descrizione dei dati in analisi	103
4.2	Addestramento reti neurali	104
4.3	Stima delle funzioni di probabilità	108
4.4	Stima dei parametri per il test del plesso corioideo	109
4.5	Stima del coefficiente k per l'identificazione della translucenza nucale	111
4.6	Risultati	112
4.6.1	Sezione sagittale mediana	112
4.6.2	Individuazione della translucenza nucale	114
4.6.3	Misurazione della translucenza nucale	115
	Bibliografia	119

Introduzione

L'evoluzione della strumentazione per l'imaging ecografico in ambiente prenatale ha permesso di approfondire sia le correlazioni che intercorrono tra marker e patologie legate all'aspetto morfologico e funzionale degli organi sia le patologie cromosomiche. Prima dell'introduzione dell'imaging ecografico neonatale, le indagini non invasive di routine per il monitoraggio del feto erano talmente limitate da non permettere di valutare l'aspetto morfologico del feto. L'uso delle attuali strumentazioni ha reso possibile studiare gli organi fetali, non solo da un punto di vista morfologico ma anche funzionale, come nel caso del cuore durante le prime settimane di gravidanza. L'analisi e lo studio di marker ecografici fetali ha permesso di individuare le correlazioni con alcune anomalie genetiche come ad esempio la sindrome di Down, causa cromosomica di ritardo mentale più diffusa (si manifesta circa in un caso ogni settecento nascite, senza distinzioni etniche o sessuali). Studi medici condotti in ambito prenatale hanno dimostrato la correlazione tra alcuni fattori come lo spessore della translucenza nucale, l'età materna e il rischio che il nascituro sia affetto da tale sindrome. La translucenza nucale (NT), sebbene non siano ancora note le cause della presenza di tale sostanza, è una falda liquida che si manifesta tra l'undicesima e la tredicesima settimana di gestazione, visibile ecograficamente nella regione nucale del feto. Negli ultimi anni, la translucenza nucale è divenuta uno dei marker ecografici di fondamentale importanza nella diagnosi prenatale di alcune patologie tra cui la trisomia 21. Inoltre lo studio combinato con altri marker ecografici, come l'ipoplasia dell'osso nasale e marker biologici presenti nel sangue materno, aumentano il livello di accuratezza delle diagnosi. Tuttavia, l'uso di ultrasuoni nella diagnostica medica è fortemente dipendente dalle competenze e dall'esperienza dell'operatore e da come questo le interpreta. Nel caso specifico tali caratteristiche possono influenzare la misura della NT. L'assenza di strumenti automatici accurati, e quindi privi della non oggettività dell'operatore,

nell'analisi delle immagini ecografiche fetali non è legato soltanto al problema della misurazione della translucenza nucale, ma anche all'individuazione della sezione sagittale mediana, requisito necessario affinché la misurazione sia attendibile. L'individuazione della sezione sagittale mediana in letteratura è stata principalmente affrontata per le acquisizioni volumetriche e sono pochi gli studi condotti in ambito bidimensionale. L'attività ricerca svolta durante il periodo di dottorato è stata orientata in tale ambito. In particolare le attività di ricerca sono state rivolte allo studio di metodologie innovative e ad una loro implementazione e ciò al fine di realizzare strumenti in grado di elaborare immagini di sequenze ecografiche fetali, acquisite in modalità bidimensionale, per estrarre da esse caratteristiche o individuare su di esse possibili patologie. La ricerca ha coinvolto attività sia di analisi di immagini a basso livello come la segmentazione basata sulle applicazioni delle trasformate wavelet sia di medio livello per il riconoscimento dei pattern con il supporto di metodi di cluster che di alto livello per l'analisi e la classificazione dei dati estratti dall'applicazione di metodiche basate sul machine learning. Le procedure che permettono l'analisi delle immagini ecografiche in modo automatico sono piuttosto articolate e presentano un'elevato grado di complessità. Il processo per individuare una buona segmentazione può influenzare positivamente o negativamente il successo dell'intero processo di analisi. In ambito medico, il processo per la segmentazione nelle immagini ad ultrasuoni risulta un problema non banale; inoltre risulta talmente di fondamentale importanza che, in molti casi clinici, esso è svolto manualmente dal medico esperto, che ne delimita, con appositi strumenti software, il contorno delle regioni di interesse.

Nel complesso le attività svolte hanno mostrato risultati promettenti, l'implementazione delle metodiche proposte ha evidenziato un loro utilizzo per supportare la diagnosi del medico, riducendo altresì l'intervento umano e quindi la dipendenza dall'operatore.

La tesi è articolata in quattro capitoli. Nel primo capitolo sono definite le basi per lo studio delle onde a ultrasuoni con alcune loro proprietà fisiche; vengono di seguito presentate alcune strumentazioni di imaging, con applicazioni ecografiche nel periodo di gestazione; si concluderà il capitolo focalizzando l'attenzione sulla translucenza nucale e il legame che essa ha sia con sindrome di Down sia con altre patologie cromosomiche. Inoltre saranno evidenziati i requisiti necessari affinché la misurazione della translucenza nucale possa essere eseguita correttamente.

Nel secondo capitolo sono introdotti i fondamenti teorici applicati durante lo sviluppo di questa tesi; viene brevemente descritta la teoria dell'analisi di Fourier, la trasformata wavelet e l'algoritmo *à trous*, attraverso i quali è stato possibile ottenere una buona segmentazione delle immagini in esame. Al fine di classificare correttamente i pattern sono state valutate le reti neurali, di cui si riportano le basi teoriche focalizzando l'attenzione sulle reti multistrato delle quali si è ne propone una versione adattata alle esigenze ecografiche prenatali. Inoltre, è descritta la Trasformata Simmetria Radiale, utile per l'individuazione rapida di strutture sufficientemente circolari.

Nel terzo capitolo viene presentata e analizzata la metodologia proposta per l'analisi della traslucenza nucale. Sono inizialmente presentati i moduli funzionali per l'individuazione della sezione sagittale mediana, specificando come l'osso mascellare e il plesso corioideo sono identificati ed analizzati. Successivamente sono presentati i moduli funzionali per l'individuazione automatica della regione nucale e per la misurazione della relativa traslucenza.

Nel quarto capitolo sono definiti i dati in analisi e sono riportate le sperimentazioni condotte; nella prima parte si esplicitano le procedure di training eseguite sulle reti neurali per l'analisi dell'osso mascellare, riportando le matrici di efficienza ottenute. In questa sezione è descritto il procedimento di campionatura per il calcolo delle funzioni di distribuzione di probabilità e la stima dei parametri utilizzati nell'analisi del plesso corioideo. Nella seconda parte vengono riportati i risultati ottenuti per i moduli funzionali per l'individuazione della sezione sagittale mediana e per la misurazione della traslucenza, infine il capitolo termina con la valutazione dei risultati e le relative conclusioni.

Capitolo 1

Ultrasuoni ed ecografia prenatale

Gli ultrasuoni sono onde acustiche poste nello spettro delle frequenze compreso tra 20kHz e 200MHz quindi frequenze superiori alla banda udibile dall'uomo; considerata l'assenza di radiazioni ionizzanti, di radiofrequenze e campi magnetici, gli ultrasuoni sono definiti non nocivi per gli organismi viventi. Questo è uno dei principali motivi per cui il loro utilizzo ha permesso alla ricerca medica di progredire, in particolare nella diagnostica e nella prevenzione in cui sono adottate frequenze dell'ordine del MHz.

L'ecografia è una tecnica diagnostica di immagine basata sull'uso di ultrasuoni, la quale permette di osservare aree esteriormente non visibili. L'ecografia fa parte di quelle tecniche diagnostiche non invasive, non traumatiche, poichè preserva l'integrità strutturale e funzionale degli organi coinvolti. Negli ultimi decenni, sono stati diversi i fattori che hanno permesso un considerevole successo degli ultrasuoni e dell'ecografia in campo medico. Primo tra tutti il loro essere non nocivi per la salute umana e l'ambiente, in quanto la generazione di tali onde, non provoca la produzione di sostanze tossiche o radioattive nocive per l'uomo, come nel caso di tomografie assiali computerizzate (TAC) e radiografie a raggi X con o senza mezzo di contrasto. Inoltre, il costo relativamente basso e la portabilità dell'apparecchiatura per l'imaging ecografico hanno contribuito ad una larga diffusione non soltanto nella sanità pubblica ma anche in quella privata.

Nonostante i numerosi vantaggi, l'uso degli ultrasuoni nella diagnostica medica presenta anche alcuni svantaggi, legati alla definizione e alla qualità delle immagini prodotte; la capacità di penetrazione degli ultrasuoni talvolta non è sufficiente per ben rappresentare organi

e tessuti con un alto valore impendente. Inoltre la presenza di tessuti adiposi, gas, particolari patologie, il rumore strumentale (speckle noise) introducono artefatti nelle immagini, rendendo necessario l'intervento di un esperto per una loro corretta interpretazione. Sebbene gli ultrasuoni siano non invasivi, è necessario rispettare le normative relative al loro impiego. Infatti, un uso improprio degli ultrasuoni può scatenare fenomeni di riscaldamento e di cavitazione (formazione di microbolle nel citoplasma). Questi effetti collaterali sono causati da intensità elevate (da 1 a 1000 W/cm²) che comportano degenerazione proteica e alterazioni del DNA e cavitazioni permanenti con conseguenti lesioni delle membrane e necrosi cellulare. Gli ultrasuoni usati in diagnostica hanno un'intensità inferiore a 1 W/cm², tale caratteristica li rende non nocivi all'essere umano. Risultati di laboratorio mostrano che tali effetti risultano dannosi solo per lunghe esposizioni a queste onde ed a frequenze inadeguate. Pertanto viene raccomandato di non usare impropriamente tali tecniche diagnostiche e di impiegarli per il tempo strettamente necessario.

Pertanto le immagini a ultrasuoni risultano un candidato ideale per la diagnostica medica e in particolare per diagnosi cliniche o di routine effettuate ad esempio durante il periodo di gravidanza. Tuttavia, i dati acquisiti da un sistema di imaging ad ultrasuoni sono di difficile interpretazione e trattazione, non solo per l'osservatore esperto che le acquisisce e le analizza, ma anche per coloro che operano nella definizione di metodi per l'analisi automatica dei dati acquisiti.

In questo capitolo verranno esposti alcuni richiami sui più importanti principi fisici legati alle onde acustiche, su cui si basano le strumentazioni di acquisizione, senza alcuna pretesa di essere un riferimento esaustivo.

1.1 Gli ultrasuoni

Gli ultrasuoni sono un sotto insieme delle onde acustiche, le quali sono onde meccaniche e possono essere prodotte da strumenti che posseggono proprietà di elasticità, inerzia e necessitano di un mezzo per propagarsi. Inoltre, portano con se energia che si propaga lungo la medesima direzione delle onde stesse. Proprietà come pressione, velocità, direzione e intensità sono necessarie per la descrizione delle onde acustiche; similmente ad altri tipi di onde, anche quelle acustiche sono soggette a fenomeni di diffrazione, riflessione e interferenza ma

non a polarizzazione in quanto l'oscillazione avviene nella stessa direzione dello spostamento. È possibile rappresentare un'onda attraverso un grafico che indichi lo spostamento effettuato da una particella durante l'oscillazione attraverso il tempo e lo spazio (Figura 1.1).

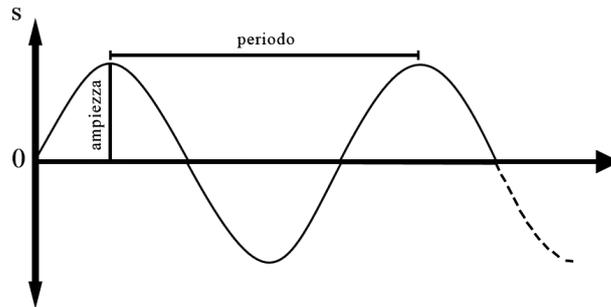


Figura 1.1: Il periodo rappresenta la lunghezza del segmento di una oscillazione completa compiuta dalla particella, mentre l'ampiezza descrive la distanza massima che compie la particella a partire dalla sua condizione di riposo. La lunghezza d'onda è la lunghezza del percorso in una oscillazione.

La velocità di propagazione del suono dipende dalle caratteristiche del mezzo in cui si propaga, direttamente proporzionale all'elasticità ed è inversamente proporzionale alla densità del mezzo, in accordo alla seguente relazione:

$$c = \sqrt{\frac{E}{\rho}}$$

Dove c indica la velocità del suono nel mezzo considerato, E indica il modulo di elasticità di Young espressa in N/m^2 e ρ la densità del mezzo espressa in Kg/m^3 . La frequenza f , misurata in Hz, rappresenta numericamente l'inverso della lunghezza d'onda L ed è proporzionale alla velocità del suono nel mezzo in cui si propaga. Il periodo T , la lunghezza d'onda e la frequenza sono legate tramite le relazioni: $f = c/L$ e $f = 1/T$.

In particolare, la frequenza in relazione alla lunghezza d'onda permette di descrivere la seguente equazione:

$$y(x, t) = y_0 \sin \omega \left(\frac{t - x}{c} \right)$$

che rappresenta la posizione della particella in un'onda acustica dove x è la distanza dalla sorgente acustica, t è l'istante di tempo, y_0 è l'ampiezza iniziale dell'oscillazione, c è la velocità dell'onda e ω è la velocità angolare.

L'ampiezza di una onda acustica è una grandezza logaritmica che misura il rapporto tra la pressione acustica P indotta dall'onda e un valore di riferimento R . Tipicamente, l'ampiezza è misurata in decibel (dB):

$$A = 20 \log \left(\frac{P}{R} \right)$$

La frequenza è considerata come fattore discriminante tra tre principali classi di onde acustiche (Figura 1.2): gli infrasuoni, il suono udibile e gli ultrasuoni. I valori 20Hz e 20kHz sono i valori che delimitano il suono udibile.



Figura 1.2: Classificazione delle onde acustiche.

Gli ultrasuoni, come tutte le onde meccaniche, hanno bisogno di una energia di input e un mezzo attraverso il quale viaggiare (solido, liquido o plasma). Una volta che l'energia è stata prodotta, le onde viaggiano attraverso il mezzo di trasmissione, portando con se l'energia. Parte di essa continuerà a propagarsi, parte verrà riflessa: la quantità di energia trasmessa e riflessa dipende dalle caratteristiche intrinseche delle strutture incontrate (i tessuti anatomici nel caso specifico del medical imaging).

Poiché la capacità di penetrazione degli ultrasuoni nei tessuti è inversamente proporzionale alla frequenza, l'uso di alte frequenze permette di visualizzare con una alta risoluzione spaziale le strutture in prossimità della sonda. La profondità dei tessuti raggiunti è dipendente dall'intervallo di tempo tra l'emissione di due impulsi consecutivi e la frequenza del segnale. Inoltre, è necessario definire il termine riflessione che è definito sul segnale di ritorno, mentre il termine scattering viene usato per indicare la dispersione del segnale riflesso.

L'interazione degli ultrasuoni con i tessuti biologici produce assorbimento dell'energia del fascio che dipende dal tipo di tessuto attraversato e dalla frequenza delle onde. In generale,

la potenza, che rappresenta l'energia trasportata dagli ultrasuoni, è generalmente minore a 1 Watt per centimetro quadrato, per evitare fenomeni di surriscaldamento e cavitazione dei tessuti trattati.

1.1.1 I sistemi ecografici

Per monitorare l'emissione degli impulsi e l'acquisizione del segnale riflesso vengono utilizzati speciali cristalli piezoelettrici in grado di trasformare una tensione elettrica in stimoli meccanici (onde) e viceversa di trasformare le onde meccaniche in tensione elettrica. Il quarzo, il titanato di bario o di piombo, lo zirconato sono alcuni dei materiali più usati. Tali cristalli sono costituiti da numerosi "dipoli" che, in presenza di un campo elettrico, modificano l'orientamento spaziale e la dimensione del cristallo. Il loro inserimento tra due elettrodi e un'opportuna differenza di potenziale causa le vibrazioni, generando onde sonore. Viceversa, le onde sonore, intercettate dai sensori, provocano una vibrazione dei cristalli piezoelettrici con la conseguente generazione di una tensione misurabile nel circuito elettrico. Tutti i dispositivi medici che prevedono l'uso di ultrasuoni sono dotati di sonde, note anche come trasduttori, progettate in modo da usufruire di entrambe le caratteristiche dei cristalli piezoelettrici.

Gli ultrasuoni vengono generati dal sistema e irradiati nei pazienti: propagato il segnale, il sistema acquisirà ad intervalli regolari e ravvicinati il segnale riflesso trasformandolo in segnale elettrico; il valore numerico sarà ottenuto misurando il segnale elettrico convertito secondo la modalità di rappresentazione fornita dal sistema. In particolare, la visualizzazione al monitor implica che l'ampiezza delle onde riflesse corrisponde all'intensità dei pixel nell'immagine. Si avrà quindi un'immagine a onde riflesse con intensità maggiore in corrispondenza dei pixel luminosi, mentre a onde riflesse con intensità minore corrisponderanno pixel scuri, dove con il termine intensità dell'onda si intende il rapporto tra la potenza propagata dall'onda e la superficie che essa attraversa.

La posizione del pixel è strettamente correlata all'intervallo di tempo intercorso tra l'emissione dell'impulso e la ricezione del segnale riflesso. In questo modo la posizione di un pixel codificherà la profondità di penetrazione dell'impulso. Ciò implica che i pixel posizionati nella parte inferiore dell'immagine corrispondono ad una profondità maggiore, mentre i pixel posizionati nella parte superiore corrispondono a regioni prossime alla superficie.

1.1.2 Sistemi di acquisizione

Gli ecografi possono essere schematizzati in tre grandi componenti: una sonda sorgente/ricevente di echi, alcuni componenti hardware per l'elaborazione del segnale e un dispositivo per la visualizzazione. Sono diverse le tipologie di strumenti che è possibile trovare in commercio in grado di emettere ultrasuoni e allo stesso tempo campionarli, e quindi in grado di misurare il segnale riflesso e rappresentarlo su un dispositivo di output. In generale nel sistema è presente una componente in grado di trasformare impulsi elettrici in meccanici e viceversa; tali impulsi saranno misurati e visualizzati in un dispositivo di output, generalmente costituito da un monitor. Durante la fase di ricezione il segnale di ritorno soggetto ad impedenza ed assorbimento degli strati attraversati viene compensato al fine di essere visualizzato. Questo tipo di amplificazione prende il nome di Time Gain Compensator ed è proporzionale al tempo trascorso tra la trasmissione e la ricezione dell'eco e quindi della profondità degli strati raggiunti. Negli ultimi anni le nuove architetture hardware hanno consentito di miniaturizzare l'intero sistema rendendolo sempre più trasportabile, maneggevole e adattabile alle varie esigenze.

1.1.3 Sonde, risoluzioni spaziali

Le sonde sono dei trasduttori, costituiti da cristalli piezoelettrici, la cui struttura asimmetrica permette di generare tensione elettrica che poi sarà tradotta in impulsi meccanici. Inoltre, sono in grado di captare gli echi di riflesso e tradurli in impulsi elettrici (Figura 1.3).



Figura 1.3: Sonde Ecografiche.

La natura del segnale elettrico generato, sinusoidale o caratterizzato da brevi impulsi, fa

sì che i cristalli piezoelettrici possano generare onde continue o intermittenti. In commercio esistono due tipologie di sonde: mono cristallo e multi cristallo. Una caratteristica importante del fascio di ultrasuoni è la zona focale in cui il fascio raggiunge la massima intensità e il minor diametro. La risoluzione di un sistema ecografico dipende dal cono di apertura con cui gli ultrasuoni sono emessi, il quale influisce sulla profondità di penetrazione in cui il fascio può considerarsi parallelo. Negli ecografi si distinguono due tipi di risoluzione: assiale e laterale. La risoluzione assiale è definita come la capacità di distinguere due oggetti vicini lungo la direzione di propagazione del fascio ed è legata alla lunghezza spaziale dell'impulso (SPL), ovvero il numero di oscillazioni in un impulso moltiplicato per la lunghezza d'onda centrale. La risoluzione laterale è invece definita come la distanza tra due oggetti vicini nella direzione perpendicolarmente alla propagazione del fascio di ultrasuoni ed è pari al suo diametro (relativamente alla profondità considerata). Il fascio può essere suddiviso nella regione di Fresnel, o campo vicino, nella quale la risoluzione laterale è massima e nella regione di Fraunhofer, o campo lontano, nella quale la risoluzione laterale degrada.

1.1.4 Modalità di rappresentazione

Le modalità di rappresentazione spesso vengono classificati in quattro macro gruppi in base alle loro caratteristiche di rappresentazione del segnale di output:

A-mode: modulazione di ampiezza

È storicamente la prima forma di rappresentazione degli ultrasuoni con forma grafica monodimensionale in cui l'asse delle ascisse esprime la profondità tra la sonda e la struttura che ha generato il segnale riflesso, mentre le ordinate indicano le ampiezze degli echi. I picchi rappresentano l'intensità dell'eco stesso. Tale rappresentazione esprime le caratteristiche fisiche delle strutture attraversate dagli echi (Figura 1.4).

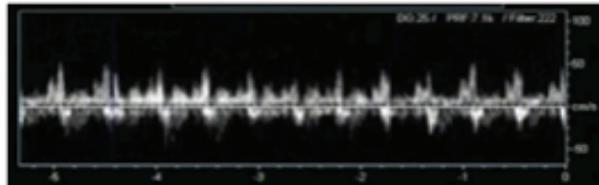


Figura 1.4: Acquisizione A-mode.

B-mode: modulazione di luminosità

L'intensità del segnale riflesso viene rappresentata con valori di grigio compresi in un intervallo che va dal nero (minima intensità) al bianco (massima intensità), proporzionali all'intensità degli echi. Ogni punto (a scala di grigio) viene rappresentato in successione, rappresentando così la distanza dalla sorgente. Le sonde in cui sono presenti diversi trasduttori permettono di avere una visualizzazione bidimensionale del segnale, ottenendo così strutture di dato bidimensionali (immagini); in questi casi la modalità è denominata B-mode Real Time e i fasci di echi vengono temporizzati e continuamente processati in almeno 15 fotogrammi per secondo. Tale modalità permette di percepire con fluidità una visualizzazione dinamica delle strutture in tempo reale, ottenendo così informazioni di tipo funzionale oltre che morfologico (Figura 1.5).

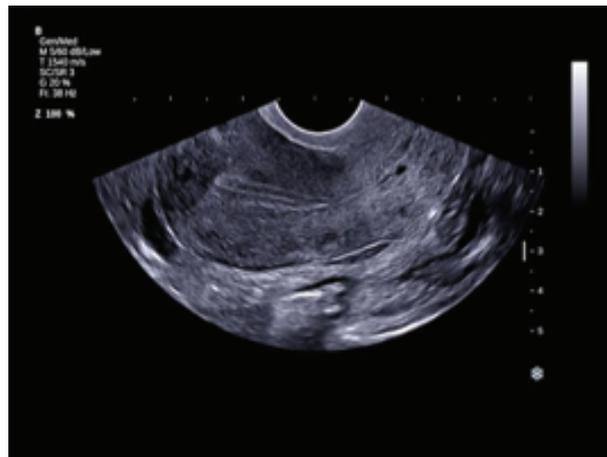


Figura 1.5: Acquisizione in B-mode.

M-mode: (motionscan)

La modalità di rappresentazione è ancora una B-mode ma con aggiornamento in tempi sufficientemente brevi degli echi riflessi; questi vengono visualizzati, sempre con punti luminosi a valori in scala di grigio, in modo consecutivo rapido. Tale modalità garantisce di visualizzare la posizione variabile di un ostacolo rispetto all'eco ed è alla base degli ecocardiografici, in cui è possibile apprezzare le fasi di contrazione e rilassamento del miocardio (Figura 1.6).

3D-mode: tridimensionalità

Gli ecografi di ultima generazione costituiscono una evoluzione del B-mode in cui le sonde sono in grado di acquisire delle sezioni volumetriche degli echi riflessi dalle strutture in-

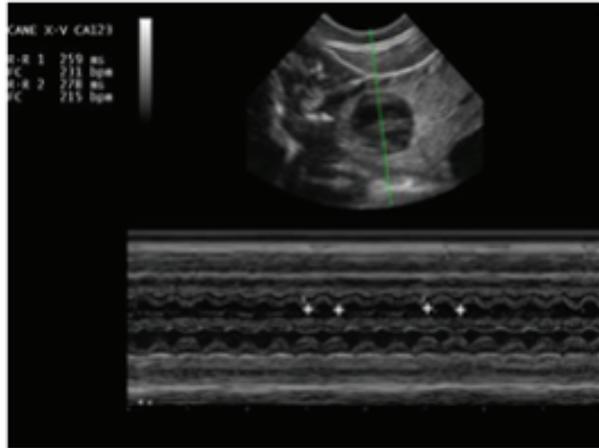


Figura 1.6: Acquisizione in Mmode

tercettate. Evoluti sistemi hardware riescono ad elaborare in brevissimo tempo tali dati e visualizzare tridimensionalmente le strutture, ruotandole attorno ai tre assi o visualizzarle in modo classico, secondo sezioni bidimensionali. Appositi software di rendering consentono di applicare delle texture ai solidi visualizzati in tempo reale e ciò al fine di mostrare le strutture in movimento e con un aspetto prossimo alla realtà (Figura 1.7).



Figura 1.7: Acquisizione in 3D-mode

1.2 Medical imaging mediante ultrasuoni

Col termine medical imaging si intende un ramo delle attività scientifiche che si occupa di studiare e analizzare processi orientati alla elaborazione delle immagini mediche. Nello specifico, le immagini ecografiche sono una rappresentazione del segnale riflesso dalle strutture, degli organi e dei tessuti, colpiti dal fascio di ultrasuoni.

In genere l'ampiezza del segnale riflesso è codificato con valori numerici e visualizzato in immagine su una scala a livelli di grigio, permettendo di ottenere in tal modo una rappresentazione pittorica dei tessuti colpiti dal segnale emesso. Nel caso dell'eco doppler, oltre ai valori a scala di grigio, sono presenti i colori, spesso utilizzati per rappresentare la dinamicità dei flussi sanguigni. Quando una sorgente di ultrasuoni emette un impulso, parte dell'energia che trasporta viene riflessa dai tessuti attraversati e parte trasmessa ai tessuti e/o organi adiacenti. Gli impulsi riflessi sono costantemente campionati su più intervalli di tempo, l'ampiezza viene misurata e la sua quantità dipenderà dall'impedenza dei tessuti attraversati. Considerando che la velocità di propagazione delle onde è sufficientemente costante, la distanza (in profondità) tra la sorgente ed il tessuto riflettente è rappresentata dal tempo trascorso tra l'emissione di un impulso e la ricezione del segnale riflesso.

1.2.1 Ecografia

L'ecografia o ecotomografia è una tecnica basata sull'uso degli ultrasuoni, largamente utilizzata in medicina sia per il monitoraggio, sia per indagini diagnostiche. Il suo scopo principale è di visualizzare le strutture interne come i tessuti e gli organi, al fine di estrarre da essi informazioni morfologiche e funzionali.

Le parti anatomiche appaiono come strutture diverse tra loro e riconoscibili in base alle loro caratteristiche di impedenza acustica e assorbimento; la presenza di eventuali anomalie nella loro rappresentazione grafica permettono allo specialista di associare la presenza di patologie come cisti, adenomi, liquido infiammatorio, degenerazioni di tessuti, formazioni tumorali benigne o maligni.

L'impedenza è una proprietà fondamentale della materia ed è una caratteristica importante per la formazione degli echi. Vi è una relazione di proporzionalità tra l'impedenza e la densità della materia attraversata alla velocità del suono. Un ruolo importante è assun-

to dall'analisi delle transizione tra due tessuti con impedenza acustica diversa denominata interfaccia che pertanto indica la zona di passaggio tra un tessuto e il suo adiacente. In corrispondenza di tali zone gli ultrasuoni vengono in parte riflessi e in parte rifratti nelle regioni con diverse caratteristiche fisiche, modificando la lunghezza d'onda e la direzione di propagazione. Le piccole aree di discontinuità, con dimensioni inferiori alla lunghezza d'onda, saranno sede di fenomeni di scattering (diffusione) in cui gli echi risultano riflessi in ogni direzione.

L'ecografo talvolta può mettere in risalto alcune elementi che la tomografia computerizzata non è in grado di evidenziare, come per esempio la steatosi epatica (accumulo di grasso negli epatociti - cellule del fegato), gli ematomi da contusione (stravaso di sangue) e altri tipi di raccolte fluide o solide isolate.

Diversi sono i fattori che permettono di identificare la frequenza più opportuna per l'analisi medica. Le alte frequenze hanno un maggiore potere risolutivo nell'immagine (le strutture raggiunte saranno ben rappresentate), ma mostrano lo svantaggio di una ridotta penetrazione negli organi, mentre le basse frequenze sono preferibili per raggiungere strutture più profonde. Le caratteristiche dei tessuti che compongono la regione in esame giocano un ruolo importante nella qualità dell'immagini acquisite e pertanto la scelta della frequenza da utilizzare deve tener conto anche dalla composizione degli strati che ci si appresta ad analizzare, poiché ciascun tessuto è caratterizzato dalla propria impedenza acustica. In genere sono suggeriti intervalli di frequenze, poiché la scelta della frequenza corretta si basa anche sulla fisiologia dell'individuo in esame: per soggetti esili si potranno utilizzare frequenze più elevate, contrariamente ai soggetti in sovrappeso, in cui lo strato adiposo sarà in un percentuale maggiore e causerà una maggiore attenuazione.

In diagnostica, a seconda del tipo sonda utilizzata, il range di frequenze varia dai 2 ai 13 MHz, mentre l'intensità nel tempo in media è dell'ordine di decine di mW/cm^2 . Inoltre, al fine di facilitare la penetrazione degli ultrasuoni nei tessuti esterni riducendo l'area tra sonda e cute esterna, viene applicato un apposito gel. All'atto del posizionamento della sonda ecografica si ricercano nel monitor i marker anatomici che permettono di orientare correttamente il piano topografico per ottenere le sezioni desiderate. Le ampiezze degli echi di ritorno verranno visualizzati tramite scala di grigio esprimendo l'ecogenicità (capacità di riflessione) del tessuto in esame. I termini iperecogeno, ipoecogeno, isoecogeno e anecogeno esprimono le proprietà

ecogeniche dei tessuti non solo in base alle loro proprietà fisiche ma anche in relazione ai tessuti circostanti. Nel caso dei tessuti anecogeni si è in presenza di tessuti poco riflettenti (nel caso di liquidi, ad esempio, l'acqua e il sangue). I valori di grigio dei tessuti anecogeni tenderanno al nero. Nel caso di solidi siamo in presenza di tessuti ecogeni in grado di riflettere il fascio in modo più o meno marcato e le relative zone nell'immagine risultano più luminose; il valore di ecogenicità del tessuto non dipenderà soltanto dalle caratteristiche fisiche del tessuto, ma anche dalla differenza delle capacità riflettenti dei tessuti adiacenti e circostanti (Figura 1.8). Il tessuto iperecogeno per eccellenza è quello osseo, che ha una bassa capacità di assorbimento riflettendo quasi interamente le onde e impedendo una corretta analisi. Esistono inoltre tessuti di natura non ossea (ad esempio, microcalcificazioni) che sono considerati iperecogeni e che, a differenza del parenchima circostante ricco di vasi sanguigni, hanno una capacità riflettente maggiore, mostrandosi quindi come regioni iperecogene. Analogamente nei tessuti ipoecogeni, anche detti a ecogenicità ridotta, e nei tessuti isoecogeni, come sinonimo di stessa ecogenicità, è difficile individuare la superficie di separazione in quanto si differenziano poco dal parenchima circostante. In definitiva l'immagine ecografica risulta una mappatura dei segnali degli echi riflessi di una sezione del corpo.

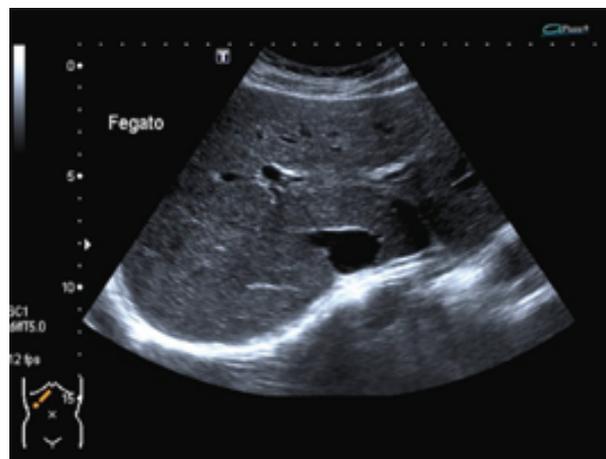


Figura 1.8: Tessuti con differente ecogenicità.

1.2.2 Artefatti

L'ecografia è, in ogni caso, una procedura fortemente dipendente dall'operatore poiché vengono richieste particolari doti di manualità e spirito di osservazione, oltre a specifiche com-

petenze sull'interpretazione delle immagini ed esperienza clinica. È proprio l'esperienza dell'operatore che permette di identificare correttamente alcuni artefatti presenti a volte sulle immagini ecografiche. Viceversa, è stato osservato che la presenza di alcuni artefatti possono esaltare alcune caratteristiche dei tessuti, aiutando così l'esperto nell'interpretazione delle immagini: la presenza di alcuni artefatti è direttamente correlata ad alcune patologie.

Tra gli artefatti più comuni si riportano:

Rifrazione: il cambiamento di velocità di propagazione determina una variazione nella direzione dell'onda, provocando un angolo di incidenza nella superficie di separazione non perpendicolare. Conseguentemente l'angolo di trasmissione risulta diverso da quello di incidenza. Ciò può determinare nell'immagine un errato posizionamento di alcune parti anatomiche o che alcune regioni non siano del tutto evidenziate (Figura 1.9).



Figura 1.9: Artefatto: “rifrazione”.

La velocità di propagazione: una errata impostazione della velocità comporta un calcolo sbagliato nella associazione tra il tempo di ritorno dell'eco e la posizione raggiunta dall'onda incidente (Figura 1.10).

Ambiguità: quando la frequenza non è impostata correttamente può avvenire che venga ricevuto un eco di un impulso precedente, originato a una diversa profondità.

Side lobes: oltre al fascio principale di ultrasuoni prodotto dalla sonda possono essere presenti dei fasci minori per intensità, disposti lateralmente e noti come lobi laterali. La loro presenza può determinare la comparsa di echi in una raccolta liquida poiché, mentre il lobo principale la attraversa senza che si creino echi, quelli laterali interagiscono con strutture o

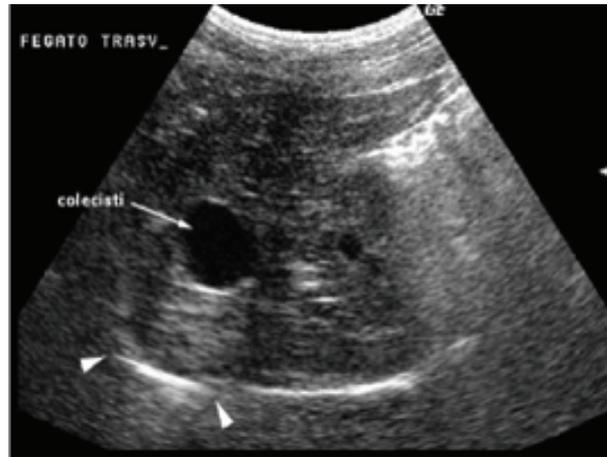


Figura 1.10: Artefatto: “velocità di propagazione”.

organi adiacenti alla raccolta e gli echi da essi generati, giunti alla sonda, vengono interpretati come facenti parte del fascio principale e disegnati nel lume della raccolta liquida. I deboli echi dei fasci laterali non creano effetti visibili in un'immagine ecogena, ma risultano visibili se buona parte dell'immagine è anecogena (Figura 1.11).

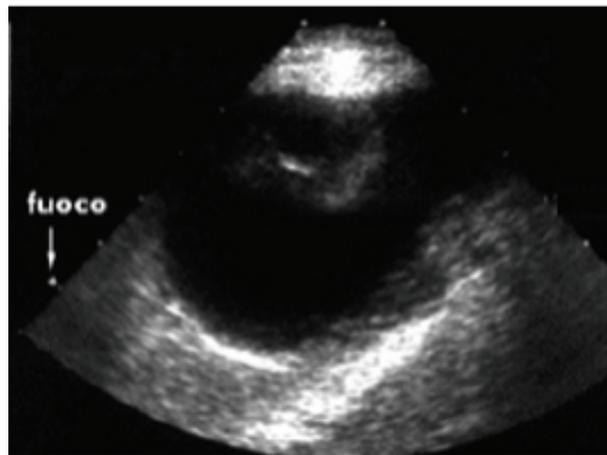


Figura 1.11: Artefatto: “side lobes”.

Immagine specchio: può altresì succedere che alcune onde vengano riflesse avanti e indietro fra la struttura e il trasduttore, provocando una doppia immagine di uno stesso organo, riflesso rispetto all'organo in cui avviene l'incidenza (Figura 1.12).

Ombre: tale artefatto si presenta quando il coefficiente di attenuazione dell'oggetto attraversato dall'onda è elevato e tutto ciò che si trova al di sotto dell'oggetto con alto coefficiente

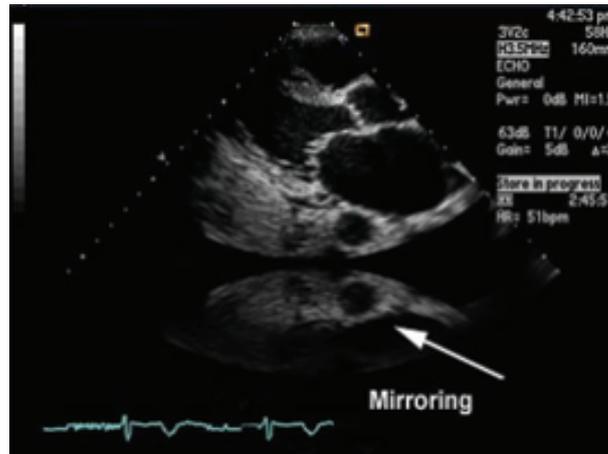


Figura 1.12: Artefatto: “specchio”.

sarà visualizzato con una bassa intensità. Questo artefatto può essere d’aiuto quando si è in presenza di sostanze con coefficiente di attenuazione anomalo come per esempio nel caso di una lesione cistica rispetto a un nodulo solido ipoecogeno (Figura 1.13).

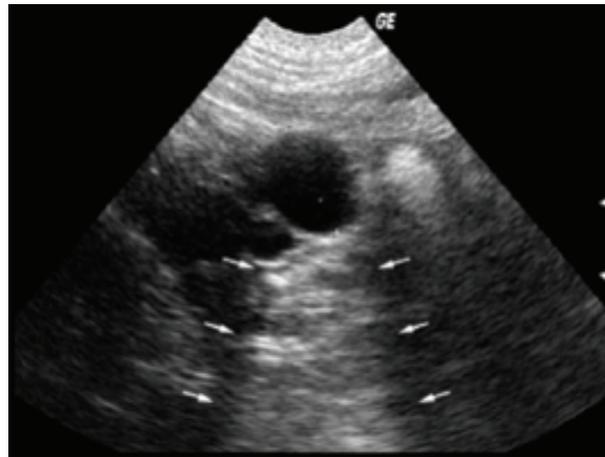


Figura 1.13: Artefatto: “Ombra”.

Riverbero: in presenza di alcune sostanze come i gas si possono presentare forti riflessioni con un rimbalzo continuo tra oggetto e trasduttore determinando una serie di bande ecogene distanziate tra loro da un intervallo costante e intensità decrescente (Figura 1.14).

Ring down: è molto simile al riverbero e si manifesta in presenza di sacche d’aria in cui il rimbalzo avviene tra la parete posteriore e anteriore creando nell’immagine un effetto a coda di cometa (Figura 1.15).



Figura 1.14: Artefatto: “riverbero”.



Figura 1.15: Artefatto “ring down”.

1.2.3 Ecografia fetale

La non invasività degli ultrasuoni ha permesso di eleggere l’ecografia come la migliore tecnica diagnostica e di monitoraggio ad alta sensibilità nel settore ostetrico. In particolare è diventata una delle principali tecniche di indagine durante l’intero periodo di gestazione.

L’ecografia prenatale (Figura 1.16) resta comunque una tra le tecniche più complesse e complete, anche se di dimensioni ridotte, riguardante l’indagine medica di un essere umano. Gli ultrasuoni dovranno propagarsi inizialmente attraverso i tessuti materni, dei quali non occorrono informazioni, e successivamente attraverso quelli fetali dei quali si pretendono informazioni. Ad oggi risulta un tipo di indagine operatore-dipendente e, più che in altri campi

medici, la capacità dell'operatore influenza la qualità grafica e informativa dell'acquisizione; inoltre non è inusuale la necessità di effettuare specifiche manovre per indurre il feto ad assumere posizioni consone all'acquisizione di immagini. Come in altri settori si ricercano durante l'acquisizione dei marker specifici per un possibile e indicativo confronto con le caratteristiche standard esistenti. Sin dalle prime settimane è possibile osservare nell'ecografia prenatale la camera gestazionale e il sacco vitellino e monitorare successivamente tutte le fasi della gravidanza. Per ogni periodo gestazionale sono stati individuati diversi marker indicativi di una corretta crescita e sviluppo del feto.



Figura 1.16: Esempio di ecografia fetale.

La camera gestazionale è il primo segno ecografico visibile a partire dalla 4° o 5° settimana dall'ultima mestruazione e alla 5° settimana avrà raggiunto un diametro di 5-6 mm e con un ritmo di crescita di circa un millimetro al giorno. Durante tale periodo il sacco vitellino è un ulteriore segno ecografico ed appare come una struttura tondeggianti all'interno della camera gestazionale. Esso costituisce un importantissimo marker durante queste settimane e permette di escludere una gravidanza anembrionica, ovvero quando si è in presenza della camera gestazionale ma non è avvenuto lo sviluppo dell'embrione. A partire dalla 6° settimana è possibile osservare l'embrione circondato dalla membrana amniotica oltre che la prima attività cardiaca. Durante questo periodo la cavità amniotica raggiunge un diametro di circa un centimetro, triplicando durante la settimana successiva.

Il sistema nazionale sanitario italiano suggerisce alle gestanti di effettuare almeno tre controlli ecografici durante l'intero periodo di gravidanza. Il controllo ecografico dovrebbero

essere eseguito almeno una volta durante ciascun trimestre, in cui è solito dividere il periodo di gestazione. Si fa riferimento a specifiche settimane per effettuare i controlli ecografici poiché alcune settimane sono più importanti di altre per l'identificazione e la misurazione di particolari marker. Ad esempio, il periodo di gestazione del primo trimestre è fondamentale per la misurazione della traslucenza nucale, considerato uno dei marker fondamentali per la diagnosi di diverse patologie cromosomiche.

Ecografia del primo trimestre (11°-13° settimana)

L'ecografia del primo trimestre fornisce numerose e importanti informazioni utili per l'intero decorso della gravidanza. Oltre che per confermare il corretto impianto dell'embrione, stabilire il numero di embrioni e l'eventuale tipo di gemellarità, sarà anche possibile ricercare specifici marker importanti per l'identificazione di alcune patologie genetiche. È in questo intervallo di tempo che viene misurato lo spessore della traslucenza nucale che, insieme a una ulteriore analisi ematiche, determina il rischio di alcune malformazioni come per esempio quelle a carico dei cromosomi 21,18,13. Inoltre è una ottima opportunità di screening per le prime valutazioni sullo stato di alcuni organi del feto come il cuore (già è completo dalla decima settimana), il cervello o gli arti (Figura 1.17).



Figura 1.17: Ecografia fetale primo trimestre.

Ecografia del secondo trimestre 20°-22° settimana

L'ecografia del secondo trimestre è denominata anche "morfologica" perché l'attenzione si concentra su alcuni aspetti morfologici osservati durante il processo di valutazione di quelle gravidanze considerate "a rischio". Durante tale periodo di gestazione il feto presenta organi e arti più sviluppati, rispetto alla precedente ecografia, consentendo una analisi più dettagliata. Il cuore, a partire dalla 20° settimana, può essere analizzato e monitorato in modo completo e accurato. Vengono analizzati e misurati nel dettaglio anche il cervello e il cervelletto con la misurazione del diametro craniale, il corretto sviluppo dell'intero apparato scheletrico e dell'apparato genito-urinario. Particolare attenzione è rivolta anche alla struttura facciale del feto e all'analisi del profilo. Tali studi hanno una incidenza non indifferente sul calcolo del rischio di patologie associate all'ecografia del primo trimestre. L'attenzione durante tale periodo non ricade solo sugli organi fetali ma anche su quelli materni coinvolti nella gravidanza come il collo dell'utero. Va osservato che la conclusione del suddetto periodo è dettata anche dalle leggi in vigore in merito all'interruzione volontaria della gravidanza (Figura 1.18).

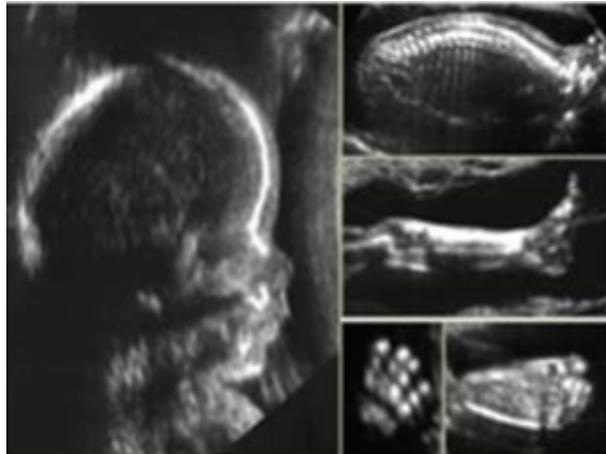


Figura 1.18: Esempio di indagine ecografica nel secondo trimestre.

Ecografia del terzo trimestre (30°-33° settimana)

Durante l'ecografia del terzo trimestre l'attenzione è rivolta a tutti quegli aspetti che possono fornire informazioni sul parto. Viene monitorato l'accrescimento del feto, la quantità di liquido amniotico e la posizione del cordone ombelicale e del feto. Sebbene la maggior par-

te delle malformazioni possono essere riscontrate con la seconda ecografia, pure durante tale periodo possono presentarsi patologie difficilmente individuabili prima di questo stadio della gravidanza come idrocefalia, tumori, ernia diaframmatica, patologie cardiache, patologie ostruttive renali (Figura 1.19).



Figura 1.19: Esempio di indagine ecografica nel terzo trimestre

1.2.4 La traslucenza nucale e le anomalie cromosomiche ad essa associata

La traslucenza nucale è una raccolta liquida sita nella regione nucale del feto che può estendersi anche lungo la colonna vertebrale. Il termine traslucenza nucale è stato coniato da Nicoladeis, pioniera nello studio prenatale della trisomia 21 alla Fetal Medicine Foundation [1, 2]. Il termine traslucenza nucale viene utilizzato indipendentemente dall'aspetto e dall'estensione della regione liquida. La traslucenza nucale ecograficamente appare come una falda scura, ovvero una regione anecogena, delimitata da due regioni iperecogene sottili (Figura 1.20). Essa è visibile sin dalle prime settimane di gravidanza e aumenta di spessore parallelamente alla crescita del feto, raggiungendo il suo massimo spessore durante l'undicesima e la tredicesima settimana; dopo tale periodo tende ad assottigliarsi. Talvolta in presenza di alcune cromosomopatie come la sindrome di Turner permane anche successivamente a tale periodo e viene denotato come "edema nucale". A partire dagli anni '90 la traslucenza nucale è diventata oggetto di una accurata analisi durante il primo trimestre di gravidanza, dopo aver riscontrato una correlazione tra il suo spessore e l'incidenza di cromosomopatie e altre anomalie. L'associazione con le anomalie cromosomiche non è determinata dall'aspetto

della NT bensì dal suo spessore: maggiore è lo spessore della traslucenza, maggiore sarà probabilità che il feto presenti alcune anomalie (Figura 1.21). Le cromosomopatie correlate ad uno spessore della traslucenza superiori ai valori di riferimento riguardano principalmente la sindrome di Down (nota anche come trisomia 21), la sindrome di Edwards (o trisomia 18), la sindrome di Patau (o trisomia 13) e la sindrome di Turner (nota anche come disgenesia gonadica), ma anche altre anomalie a carico del cuore (ad esempio nella valvola tricuspide) o di altri organi come la spina bifida, l'onfalocele o l'ernia diaframmatica [3, 4]. La misurazione viene effettuata anche nel caso di gravidanze plurime, in quanto risulta ugualmente accurata e il calcolo del rischio viene effettuato per ogni singolo embrione.

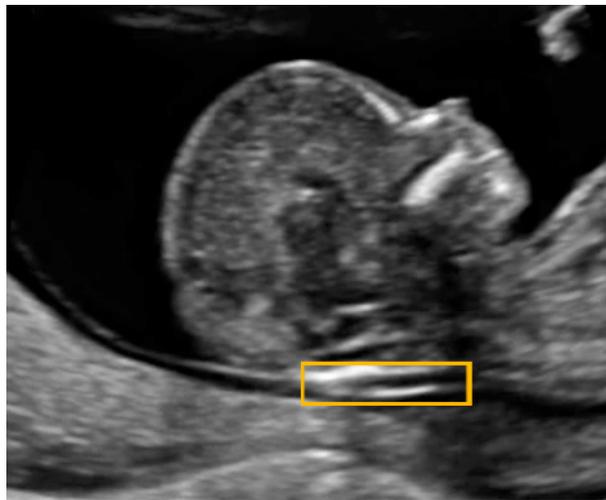


Figura 1.20: Il rettangolo evidenzia le regioni iperecogene che delimitano la traslucenza.

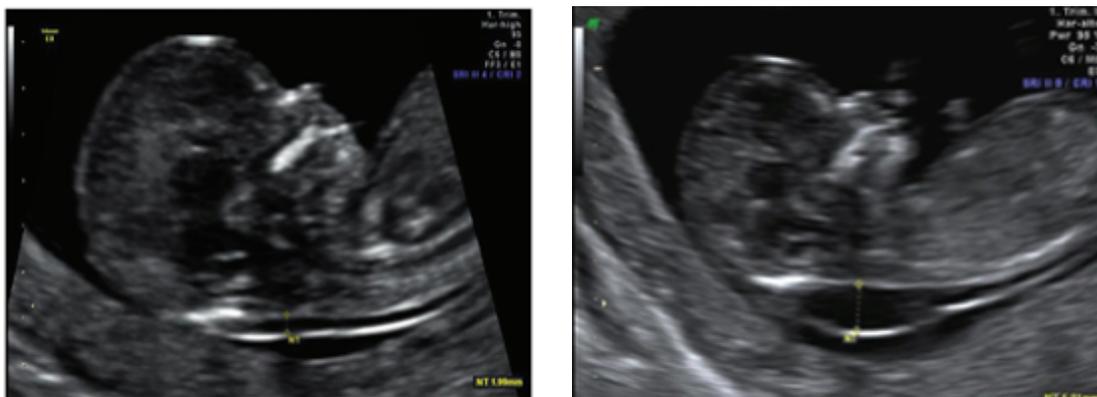


Figura 1.21: Translucenza nucale con spessore nella norma (sinistra) e superiore (destra).

Nel caso specifico della sindrome di Down e prima della scoperta della relazione con lo spessore della translucenza nucale, l'unico metodo di screening tradizionale era l'età materna. Questo test, rivelatosi insufficiente, permetteva di indicare il rischio di essere in presenza di un feto affetto da trisomia 21 con una percentuale di identificazione del 30%. Negli anni '90 studi condotti su markers biochimici presenti nel siero materno (dosaggi della Free-Beta HCG e della Pregnancy Associated Plasma Protein A (PAPP-A)), combinati all'età materna permettono di raggiungere una percentuale di corretta identificazione tra il 50% e il 70% [5]. Oggi grazie alla misurazione della translucenza nucale, combinata ai test biochimici e l'età materna, è possibile raggiungere una percentuale di identificazione del 85%-90% e ridurre il numero dei falsi positivi, cioè degli allarmi inutili, al di sotto del 5%.

A supporto dell'identificazione della sindrome di Down esistono ulteriori marker ecografici, come la misurazione dell'osso nasale associata all'età materna e alla misurazione della translucenza nucale che consentono di identificare il 90% dei feti con tale sindrome. Combinando insieme tutti questi marker si raggiunge il 95%, migliorando l'efficacia del test e riducendo il numero dei falsi positivi.

A supporto della misurazione della translucenza nucale per l'identificazione di feti affetti dalla Sindrome di Down sono stati studiati altri marker ecografici: l'osso nasale e l'osso frontale, in quanto i pazienti sono soggetti ad ipoplasia o addirittura assenza dell'osso nasale e presentano un osso frontale piuttosto piatto. In letteratura esistono diversi studi [6, 7] che evidenziano come il 70% dei feti in assenza o ipoplasia dell'osso nasale siano affetti dalla sindrome di Down.

Tra i test che è possibile effettuare durante il primo trimestre non bisogna dimenticare l'amniocentesi e la villocentesi di carattere genetico. L'amniocentesi consiste nel prelievo transaddominale del liquido amniotico dalla cavità uterina e la villocentesi (prelievo dei villi coriali) consiste nell'aspirazione di una piccola quantità di tessuto coriale. Tali metodiche diagnostiche presentano un'accuratezza maggiore, raggiungendo una percentuale di identificazione del 99% rispetto alle metodiche descritte precedentemente, ma entrambe sono definite invasive con un rischio di aborto dell'1% per la villocentesi e circa il 2% per l'amniocentesi [8, 9] e con rischio di lesioni fetali, parto prematuro, complicanze settiche, rottura prematura delle membrane e altre complicazioni.

1.2.5 La misurazione della traslucenza nucale

La Fetal Medicine Foundation è un ente non ai fini di lucro che mira a migliorare la salute delle donne in gravidanza e dei loro bambini attraverso la ricerca e la formazione in medicina fetale. La Fondazione, con il sostegno di un gruppo internazionale di esperti, ha introdotto un programma educativo sia per i professionisti sanitari sia per i genitori, dando la possibilità di ottenere certificazioni di idoneità in diversi ambiti della medicina fetale. Corsi specifici forniscono una adeguata preparazione per i medici che intendono occuparsi del monitoraggio dei feti del primo trimestre tra cui la misurazione della traslucenza nucale. Come per gli altri marker ampiamente accettati durante la gestazione, anche per la traslucenza nucale è stato redatto un protocollo per una corretta e accurata misurazione.

I criteri riconosciuti dalla Fetal Medicine Foundation per determinare lo spessore della traslucenza nucale sono [4]:

- l'età gestazionale deve essere compresa tra l'undicesima e la tredicesima settimana;
- la lunghezza vertice sacrale deve essere compresa tra i 45 mm e gli 84 mm;
- nella immagine bisogna includere soltanto la testa fetale e la porzione superiore del torace;
- visualizzare il profilo del volto fetale;
- il feto deve essere in posizione neutrale;
- l'ingrandimento deve essere il maggiore possibile;
- bisogna ottenere una buona sezione sagittale mediana del feto;
- deve essere posta attenzione nel distinguere la cute fetale dalla membrana amniotica poiché entrambe le strutture appaiono come sottili membran, a quest'epoca gestazionale;
- deve essere misurato lo spessore massimo della traslucenza sottocutanea;
- durante l'ecografia, deve essere ottenuta più di una misurazione della traslucenza nucale e la maggiore tra queste deve essere annotata.

Per ottenere risultati affidabili è necessario seguire i criteri nella misurazione della traslucenza nucale; tralasciarne alcuni comporterebbe una diagnosi sbagliata. Nel caso di nuca fetale iperestesa (Figura 1.22-*sinistra*) la misurazione può risultare aumentata di 0.6 mm. Viceversa quando la nuca è flessa (Figura 1.22-*destra*) la traslucenza nucale può essere sotto-stimata di 0.4 mm. Analogamente, misurazioni falsate si ottengono da sezioni sagittali non mediane (Figura 1.23-*destra*) in cui è possibile notare lo spessore della traslucenza visibilmente aumentato rispetto a sezioni sagittali mediane (Figura 1.23-*sinistra*).

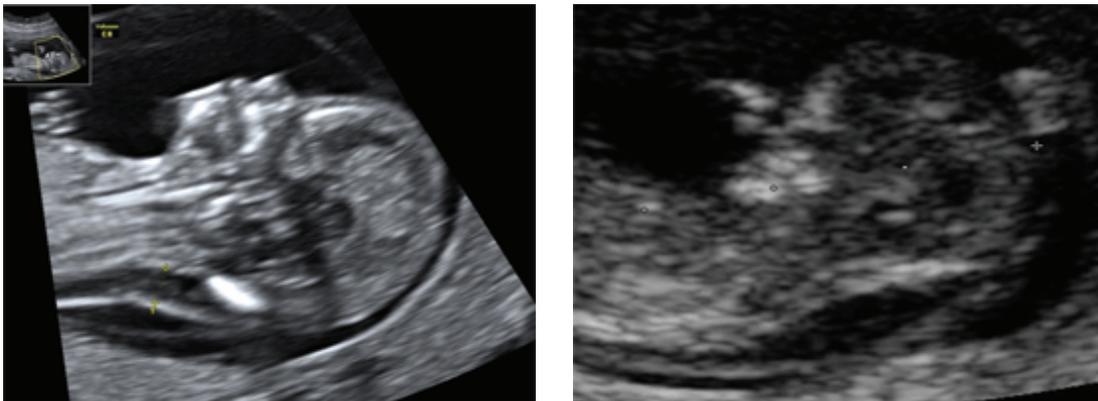


Figura 1.22: Nuca iperestesa (sinistra) e nuca flessa (destra).

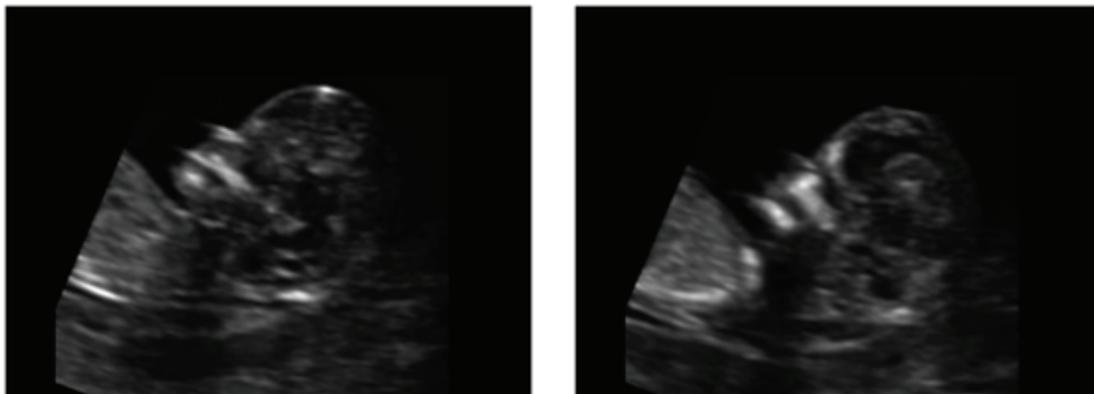


Figura 1.23: La traslucenza appare con le dimensioni corrette in a una sezione sagittale mediana (sinistra), ma presenta uno spessore maggiore in una sezione sagittale non mediana (destra).

Piani anatomici

Attraversando il corpo umano con dei piani immaginari si creano i cosiddetti piani anatomici, che si differenziano a seconda del verso e del punto di taglio. I principali piani usati in medicina per descrivere le sezioni del corpo sono tre:

- coronale o frontale;
- sagittale;
- trasverso.

Riferendoci ad un corpo in posizione eretta, il piano coronale divide dall'alto verso il basso il corpo in due parti (anteriore e posteriore) ed è parallelo alla fronte. Il piano sagittale attraversa il corpo dalla parte anteriore a quella posteriore dividendolo in due metà (destra e sinistra). Il piano trasversale è parallelo alla base del corpo e lo divide in due parti (superiore e inferiore). Si denota con sagittale mediano quel piano che divide il corpo in due metà simmetriche, attraversandolo centralmente (Figura 1.24).

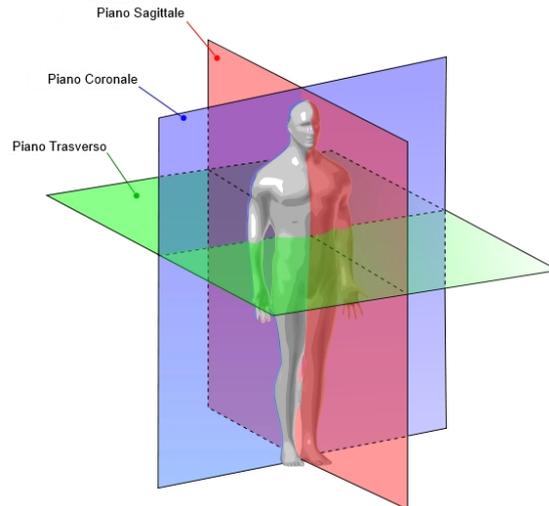


Figura 1.24: Schema dei piani anatomici.

1.2.6 Sezione sagittale mediana e traslucenza nucale

Uno dei punti formulati nel protocollo della FMF fa esplicito riferimento alla individuazione di una buona sezione sagittale. Ecograficamente si ottiene una buona sezione sagittale

mediana del cranio fetale e delle relative parti anatomiche presenti nell'immagine quando la sonda ecografica è posizionata in corrispondenza della falce cerebrale che divide i due emisferi cerebrali. Per determinare se l'immagine in esame è una sezione sagittale mediana si prendono in considerazione due elementi rappresentativi: il plesso corioideo e l'osso mascellare (Figura 1.25).

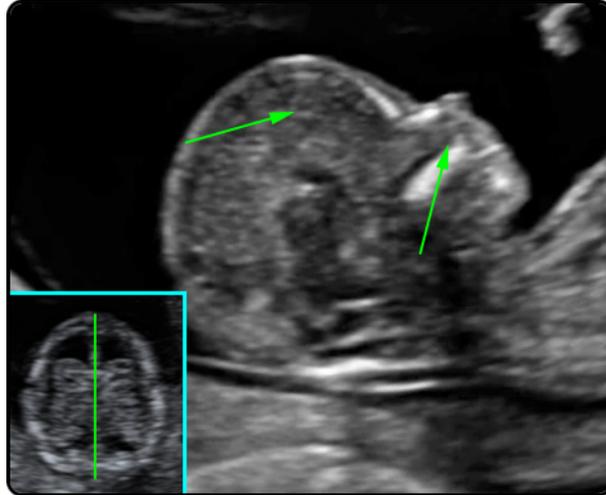


Figura 1.25: Nella sezione sagittale mediana le frecce indicano la regione ecogena del plesso e l'assenza del ramo verticale nell'osso mascellare.

Come mostrato in Figura 1.25, un buon posizionamento della sonda ecografica (in corrispondenza della falce cerebrale) permette di visualizzare una regione uniformemente ecogena nella regione del plesso corioideo. Diversamente quando la sonda non è in una posizione ecograficamente centrale verrà visualizzata una formazione tondeggianti, circondata interamente o parzialmente da una regione anecogena più o meno estesa a seconda dell'angolo di inclinazione che la sonda idealmente forma con un piano trasversale che attraversa il cervello. Il suo aspetto è dovuto alle sue caratteristiche fisiche: trattandosi di una sostanza liquida (fluido cerebrospinale) apparirà nell'immagine come una zona anecogena (Figura 1.26).

L'osso mascellare è un altro importante elemento per la determinazione della sezione sagittale mediana: se la sonda ecografica non è correttamente posizionata nell'immagine sarà visualizzato il ramo verticale dell'osso mascellare (Figura 1.26), ovvero la parte che congiunge la cavità nasale con il processo frontale (Figura 1.27).

Affinché la sezione sagittale sia definita anche mediana è necessario che non siano visibili

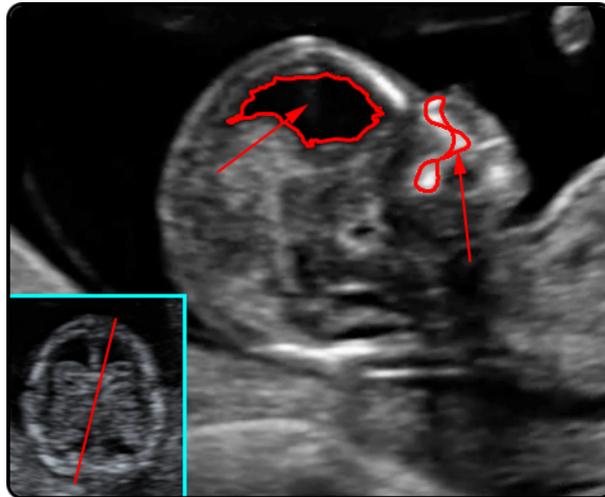


Figura 1.26: Sezione sagittale non mediana: le frecce indicano la regione anecogena del plesso e il ramo verticale dell'osso mascellare.

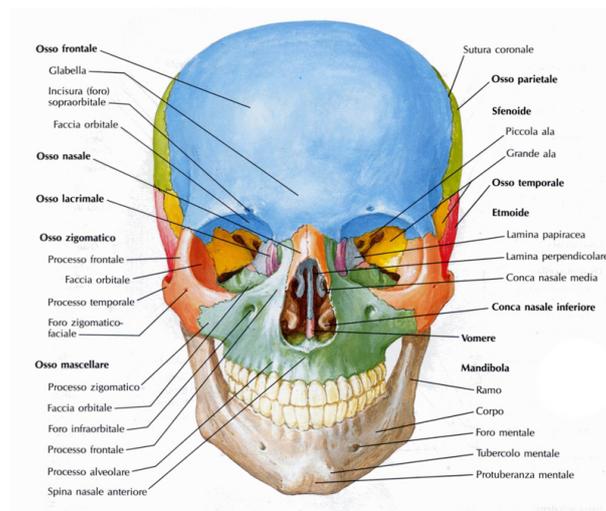


Figura 1.27: Ossa del cranio.

il ramo verticale dell'osso mascellare e che la porzione della immagine contenente il plesso corioideo appaia ecogena e priva di regioni anecogene a ridosso dell'osso frontale.

Identificata la sezione sagittale mediana, l'attenzione è focalizzata sulla translucenza nucale e sulle due regioni che la delimitano. È necessario distinguere la cute fetale dalla membrana amniotica che appaiono entrambe come sottili membrane. I caliper, come specificato dal protocollo, devono essere posizionati sulle linee che definiscono la translucenza nucale prestando

attenzione a non posizionarli all'interno del fluido nucale. In Figura 1.28 sono schematizzati i possibili casi di posizionamento corretto e non corretto dei caliper.

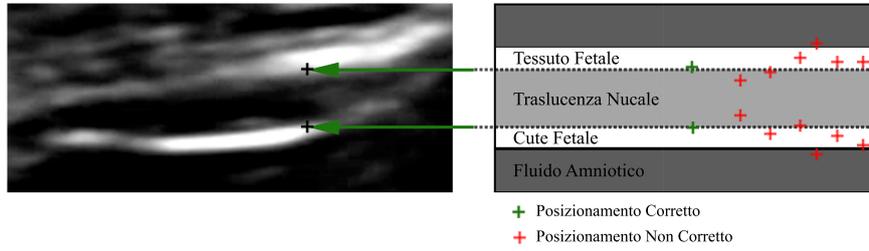


Figura 1.28: Posizionamento dei caliper per la misurazione della traslucenza nucale.

Capitolo 2

Metodi e tecniche

Ad oggi nel campo dell'elaborazione delle immagini digitali non è stata sviluppata alcuna metodologia universalmente riconosciuta in grado di comprendere e interpretare singoli fotogrammi o intere sequenze video. L'analisi automatica delle informazioni focalizza i propri interessi in diversi ambiti scientifici e tecnologici: diagnostico, sorveglianza, monitoraggio che rappresentano soltanto alcuni dei possibili ambiti di applicazione. Generalmente le fasi che compongono una specifica metodologia sono indirizzate a particolari compiti che possono essere eseguiti in modo parallelo o sequenziale. L'iter generale per una completa analisi di dati bidimensionali è costituito da operazioni volte al miglioramento delle immagini e ciò al fine di esaltare quelle caratteristiche utili a identificare ed estrapolare oggetti dal contesto per analizzarne il loro significato. Solitamente dettata dalle applicazioni che si intende sviluppare, l'attenzione è focalizzata solo su quella parte di informazione che permette di raggiungere l'obiettivo prefissato. L'identificazione dei punti di interesse è una branca dell'analisi di immagine il cui scopo è quello di delimitare regioni in cui sono racchiuse informazioni chiave. Talvolta sono necessarie elaborazioni di basso livello che permettono di agire sull'immagine di input in modo da esaltarne o sopprimerne alcune caratteristiche (ad esempio, denoising). Una completa trattazione dei metodi e delle tecniche presenti in letteratura risulterebbe complessa e dispersiva, di seguito saranno esposti soltanto gli argomenti affrontati durante il periodo di dottorato.

2.1 Trasformata wavelet

Nei prossimi paragrafi saranno riassunti alcuni importanti concetti matematici riguardanti le wavelet, partendo dalla analisi di Fourier fino ad introdurre la teoria delle wavelet, definiremo le trasformate ad esse associate e le loro principali proprietà. Per ulteriori approfondimenti si rimanda a [10, 11, 12].

La ben nota teoria formulata da Jean-Baptiste Joseph Fourier risulta estremamente efficiente in diversi campi come quello ingegneristico, fisico e matematico. La sua applicazione ha permesso di rivoluzionare la rappresentazione del segnale, scomponendolo in componenti elementari, in modo tale da metterlo in relazione con le sue proprietà fisiche. L'analisi di Fourier ben si presta alla elaborazione di segnali stazionari (costanti nel tempo) in quanto determina quali frequenze sono presenti nel segnale senza indicare quando queste si manifestano. Proprio per superare questa limitazione, il matematico Dennis Gabor sviluppò un nuovo approccio basato sulla trasformata di Fourier, utile allo studio dei segnali non-stazionari (con caratteristiche variabili nel tempo) denominata Short-Time Fourier Transform (STFT). La dimensione della finestra con cui il segnale viene spezzato e analizzato determina la precisione dell'analisi. Tale artefatto introduce del "rumore" nel segnale originario, e dipende principalmente dalla particolare forma della finestra utilizzata. Con lo sviluppo della trasformata wavelet viene introdotto il concetto di "finestra a dimensione variabile" che viene tralata sul segnale suddividendolo in porzioni sempre più piccole su cui calcolarne lo spettro. Si è introdotto così il concetto di multi-risoluzione, legato alla risoluzione delle finestre utilizzate nelle varie iterazioni. La teoria delle wavelet permette una analisi più adeguata dei segnali non stazionari.

Nel caso specifico delle immagini si individuano componenti a basse frequenze, che in genere rappresentano il background, e componenti ad alte frequenze, che rappresentano la presenza di regioni di discontinuità come i contorni. Attraverso la trasformata wavelet è dunque possibile non solo individuare le frequenze presenti nel segnale ma anche quando queste si manifestano, colmando la perdita di informazione create dalle tecniche precedenti.

Le principali proprietà osservabili sulle wavelet possono essere sintetizzate in tre punti chiave:

- permettono di analizzare segnali non stazionari, grazie alla sua capacità di localizzazione

tempo-frequenza;

- forniscono una rappresentazione multi risoluzione;
- computazionalmente sono di facile realizzazione mediante l'uso di un banco di filtri.

2.1.1 Trasformata di Fourier

La trasformata di Fourier (FT) fu sviluppata dal matematico francese Fourier nel 1822 per dimostrare che ogni funzione periodica può essere espressa come somma infinita di funzioni esponenziali complesse [13]. Dato un segnale $f(t) \in \mathbf{L}^2(\mathbb{R})$ nel dominio del tempo, la trasformata di Fourier ne dà una rappresentazione nel dominio della frequenza ν dove $\mathbf{L}^2(\mathbb{R})$ è lo spazio delle funzioni a quadrato sommabile, ovvero funzioni che godono della proprietà: $\int_{-\infty}^{+\infty} |f(x)|^2 dx < \infty$.

La FT e la sua inversa IFT sono definite dalle seguenti espressioni matematiche:

$$FT_f(\nu) \equiv \hat{f}(\nu) = \int_{-\infty}^{+\infty} f(t)e^{-j2\pi\nu t} dt \quad (2.1.1)$$

$$IFT_f(t) \equiv \hat{f}(t) = \int_{-\infty}^{+\infty} f(\nu)e^{j2\pi\nu t} d\nu \quad (2.1.2)$$

Dalla (2.1.1) si deduce la totale perdita di informazione legata alla risoluzione temporale: non è possibile determinare in quale intervallo di tempo le frequenze individuate dalla trasformata si manifestano. Analogamente, il segnale $f(t)$ non consente di determinare quali frequenze sono presenti, avendo una risoluzione nulla sull'asse delle frequenze e infinita sull'asse dei tempi. Questo è uno dei principali motivi per cui la trasformata di Fourier non risulta adatta per l'analisi di segnali variabili nel tempo. Come è possibile vedere nelle Figure 2.1 e 2.2, due segnali visibilmente diversi con le medesime componenti in frequenza, possono presentare spettri di frequenze pressoché identici poiché l'informazione legata al tempo viene nascosta. La trasformata non permette inoltre di esprimere quelle frequenze, presenti nel segnale, che variano uniformemente (Figura 2.3). Tale inconveniente nasce dal fatto che l'integrale è esteso su tutto l'intervallo temporale $] -\infty, +\infty[$, perdendo così ogni informazione circa l'istante esatto in cui una data frequenza è individuata nel segnale.

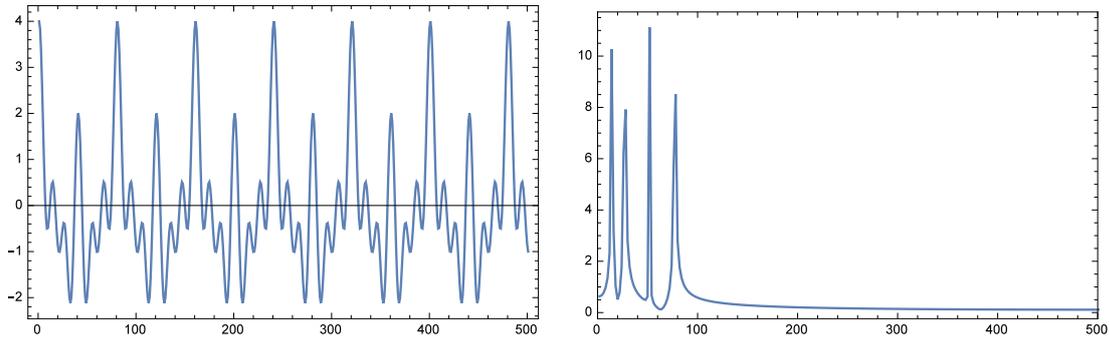


Figura 2.1: Rappresentazione nel dominio del tempo (sinistra) e delle frequenze (destra) di un segnale stazionario.

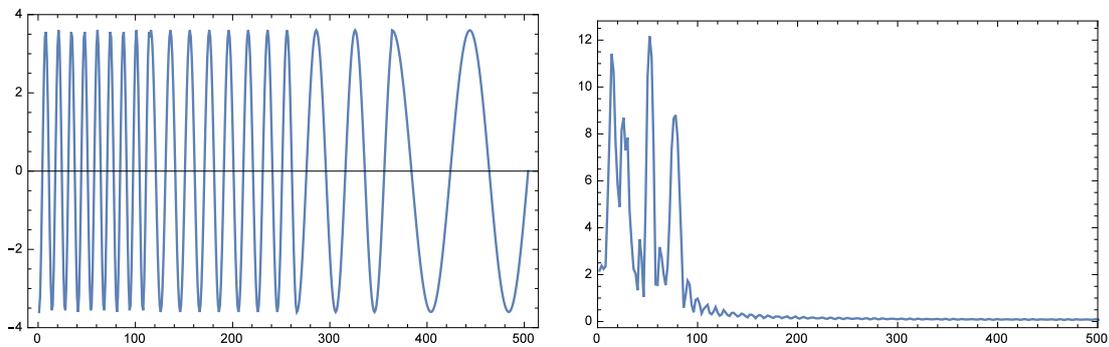


Figura 2.2: Rappresentazione nel dominio del tempo (sinistra) e delle frequenze (destra) di un segnale non stazionario.

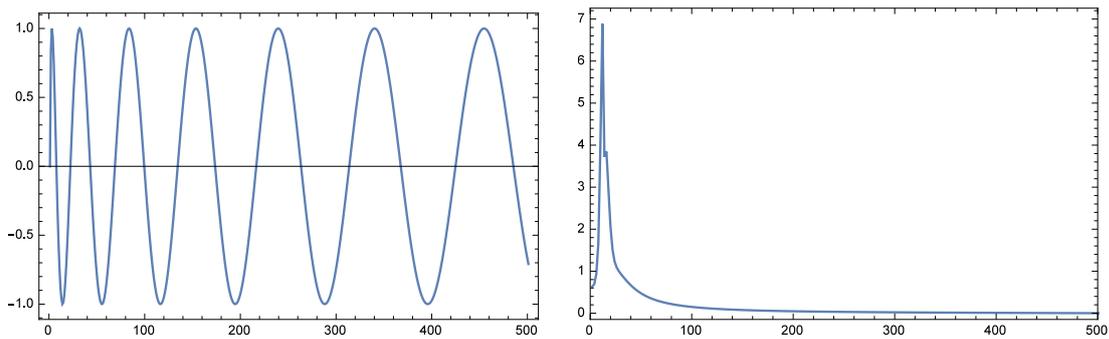


Figura 2.3: Rappresentazione nel dominio del tempo (sinistra) e delle frequenze (destra) di un segnale con cambio costante in frequenza.

2.1.2 Short Time Fourier Transform

La Short Time Fourier Transform (STFT) è stata introdotta ad opera del matematico Dennis Gabor [14] per soddisfare l'esigenza di mantenere una dipendenza della locazione temporale

delle frequenze all'interno del segnale, in STFT è stato introdotto il concetto di finestra g che permette di analizzare il segnale localmente piuttosto che nella sua interezza. La finestra $g(t - \tau)$ viene spostata sull'asse del tempo e pertanto la STFT elabora diverse posizioni τ (Figura 2.4). La STFT ha la seguente espressione matematica:

$$STFT_f(\tau, \nu) = \int_{-\infty}^{+\infty} f(t) g^*(t - \tau) e^{-j2\pi\nu t} dt \quad (2.1.3)$$

dove $*$ coincide con l'operazione di complesso coniugato e τ rappresenta il parametro di traslazione della finestra lungo l'asse del tempo t . Con la progressiva traslazione, vincolata a finestre sufficientemente piccole, su tutto il dominio della $f(t)$ si otterrà un'analisi dell'intero segnale.

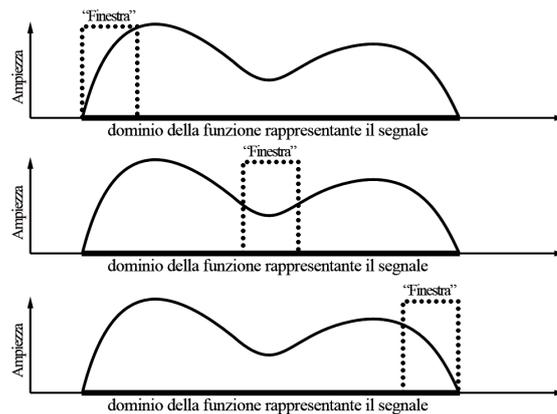


Figura 2.4: Visualizzazione del processo di STFT.

Moltiplicare il segnale $f(t)$ con g^* equivale di fatto ad effettuare la convoluzione dei loro spettri $F(\nu)$ e $G(\nu)$ nel dominio della frequenza. Il vantaggio evidente della STFT risiede nella capacità di localizzare il contenuto in frequenza in un intervallo Δt sotto il vincolo di inversa proporzionalità tra l'incertezza Δt sull'asse del tempo e quella $\Delta \nu$ sull'asse delle frequenze.

Una buona trasformata che operi nel tempo-frequenza dovrebbe avere Δt e $\Delta \nu$ molto piccoli. Tuttavia, per analogia al principio di indeterminazione di Heisenberg, si esclude la possibilità di avere una risoluzione arbitrariamente elevata sia nel tempo che in frequenza,

ossia il vincolo può essere espresso dalla seguente formula:

$$\Delta t \Delta \nu \geq \frac{1}{2}$$

Poiché la STFT non è in grado di discriminare due toni puri la cui distanza in frequenza è inferiore a $\Delta \nu$, la scelta della finestra g non è del tutto banale. Inoltre la sua forma è immutabile durante il corso dell'analisi e non permette di avere contemporaneamente una buona localizzazione sia sull'asse dei tempi sia sull'asse delle frequenze. Di conseguenza, scegliendo finestre di larghezza ridotta si ottiene un'alta risoluzione temporale, riuscendo a discriminare eventi anche prossimi tra loro. Viceversa, scegliendo finestre con una larghezza maggiore si ottiene una peggiore risoluzione lungo l'asse del tempo. Questo concetto è graficamente rappresentabile attraverso i cosiddetti “rettangoli di Heisenberg” (Figura 2.5) in cui l'area coperta da ogni rettangolo resta costante mentre varia l'incertezza lungo gli assi in modo inversamente proporzionale:

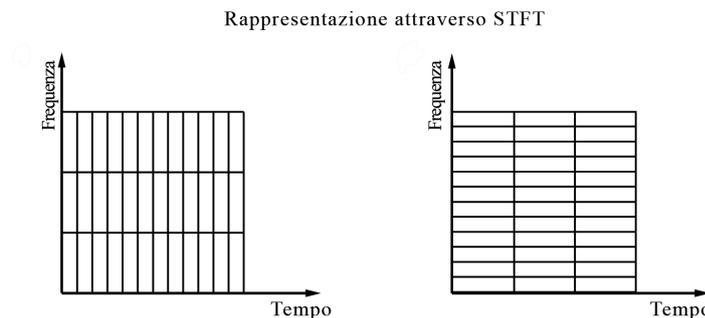


Figura 2.5: Rappresentazione dei “rettangoli di Heisenberg”.

È dunque evidente che l'immutabilità della finestra durante l'analisi del segnale comporta implicitamente una forte limitazione della STSF, rendendo necessaria un'analisi che preveda l'uso di una finestra “variabile” in modo tale da rappresentare non solo il contenuto in frequenza, ma contemporaneamente mantenere una buona localizzazione temporale.

2.1.3 La trasformata wavelet continua

La trasformata wavelet continua fu introdotta da Aleksandar Grossmann e da Jean Morlet [15] con le medesime potenzialità di rappresentazione tempo-frequenza della STFT, pur presentando un approccio differente per quanto riguarda la risoluzione con cui il segnale viene

esaminato. Essa opera con una funzione finestra la cui ampiezza varia nel corso dell'analisi (Figura 2.6), ma in questo caso le funzioni finestra di base sono generate traslando e scalando un'unica funzione non nulla ψ (nota come *wavelet madre*) a media zero $\int_{-\infty}^{+\infty} \psi(t) dt = 0$ definita in $\mathbf{L}^2(\mathbb{R})$:

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) \quad (2.1.4)$$

dove s e u sono rispettivamente il parametro di scala e di traslazione.

La trasformata wavelet di un segnale è ottenuta attraverso una correlazione:

$$W_f(u, s) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{s}} \psi^*\left(\frac{t-u}{s}\right) dt \quad (2.1.5)$$

dove l'ampiezza $W(u, s)$ indica quanto la funzione $f(t)$ assomiglia alla wavelet $\psi_{u,s}$ e il fattore di normalizzazione \sqrt{s} mantiene costante l'energia del segnale per ogni valore di scala considerato.

Applicando il teorema di Parseval la (2.1.5) può essere riscritta come integrale nel dominio delle frequenze:

$$\int_{-\infty}^{+\infty} f(t) \psi^*\left(\frac{t-u}{s}\right) dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\nu) \hat{\psi}^*(\nu) d\nu,$$

Il coefficiente $W_f(u, s)$ dipenderà dai valori assunti da $\hat{f}(\nu)$ e $f(t)$ nelle regioni tempo-frequenza in cui è concentrata l'energia di $\psi_{u,s}$ e $\hat{\psi}_{u,s}$. Quest'ultima funzione $\hat{\psi}_{u,s}$ indica la trasformata di Fourier ed è calcolata come:

$$\hat{\psi}_{u,s} = e^{-i\nu u} \sqrt{s} \hat{\psi}(s\nu).$$

Grossmann e Morlet hanno dimostrato che una perfetta ricostruzione del segnale f è garantita se la wavelet ψ soddisfa la seguente *condizione di ammissibilità*:

$$C_\psi = \int_{-\infty}^{+\infty} \frac{|\hat{\psi}(\nu)|}{\nu} d\nu < +\infty.$$

allora la trasformata $W(u, s)$ è invertibile e la sua trasformata inversa è espressa tramite la seguente relazione:

$$f(x) = \frac{1}{C_\psi} \int_{-\infty}^{+\infty} W(u, s) \psi_{u,s}(x) \frac{du ds}{2}$$

A partire dall'insieme infinito dei coefficienti della trasformata wavelet $W(u, s)$ è possibile ricostruire la funzione $f(x)$. Saranno i parametri u e s a determinare la risoluzione temporale e frequenza. Ad alti valori del parametro s corrisponderà un'alta risoluzione in frequenza e bassa risoluzione temporale; sarà quindi possibile individuare eventi di breve durata in cui il segnale presenta delle alte frequenze. Viceversa, per bassi valori di s si avrà una bassa risoluzione in frequenza e un'alta risoluzione temporale; individuando eventi di lunga durata a bassa frequenza. Il parametro u indica l'istante in cui si verifica una determinata frequenza.

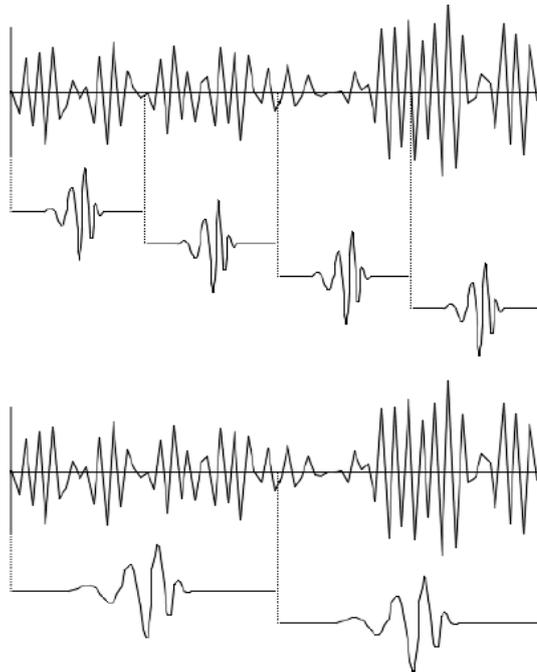


Figura 2.6: Processo di scala e traslazione.

Intuitivamente, ciò significa che il piano tempo-frequenza rappresentato attraverso i rettangoli di Heisenberg sarà tassellato con rettangoli aventi la stessa area con base e altezza con dimensioni proporzionali (Figura 2.7).

Il fatto che la trasformata wavelet sia una rappresentazione bidimensionale di un segnale monodimensionale indica l'esistenza di una ridondanza nei suoi coefficienti. Nel caso di

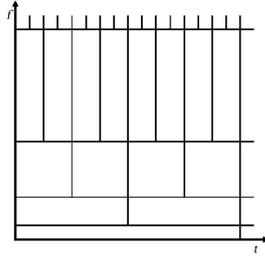


Figura 2.7: Rappresentazione del piano tempo-frequenza della trasformata wavelet.

segnale discreto, è possibile effettuare l'analisi su un numero finito di valori per u ed s , ovvero campionando i parametri di scala e traslazione. Inoltre per ogni coppia di valori u ed s , tale metodo deve approssimare numericamente l'integrale del prodotto tra il segnale e la wavelet $\psi_{(u,s)}$. In tal modo l'implementazione numerica dell'analisi wavelet continua su un segnale discreto produce un numero di coefficienti pari al prodotto tra il numero di valori differenti assunti da u ed il numero di valori differenti assunti da s .

Mallat nel 1986 ha formalizzato un algoritmo per l'analisi wavelet, che consente una analisi computazionalmente più efficiente ma, allo stesso tempo, egualmente esauriente. Mallat ha proposto di discretizzare l'insiemi dei valori finiti assunti dai parametri u ed s così da determinare una rappresentazione del segnale completa e non ridondante [12].

La possibilità di ricostruire una funzione continua a partire da un insieme di campioni discreti è regolata dal teorema di Nyquist, secondo il quale la frequenza di campionamento non può essere inferiore al doppio della massima frequenza contenuta nella funzione campionata.

2.1.4 Trasformata wavelet discreta

Per ottenere una discretizzazione efficiente, al fine di ricoprire efficientemente l'asse dei tempi, è necessario tenere in considerazione che a un valore di scala piccolo dovrà corrispondere una valore di traslazione piccolo. In modo analogo per gradi valori del parametro di scala il valore del parametro di traslazione potrà essere maggiore. Una possibile discretizzazione prevede la variazione del parametro di traslazione in funzione del parametro di scala, ponendo: $s = s_0^{-j}$ e $u = ks$ con $j, k \in \mathbb{Z}$ dove s rappresenta una discretizzazione logaritmica dell'incertezza relativa a $\Delta\nu/\nu$. Nel caso in cui $s_0 = 2$ il campionamento prende il nome di diadico.

Si definisce la famiglia delle wavelet:

$$\psi_k^j(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) = 2^{j/2} \psi(2^j t - k) \quad (2.1.6)$$

dove i valori di j e k variano in un insieme finito di elementi appartenenti a \mathbb{Z} .

Mallat ha sviluppato l'analisi multirisoluzione per ricavare un algoritmo efficiente per il calcolo della DWT, introducendo concetti di funzioni di scaling e operatore di approssimazione. Quindi è possibile implementare l'analisi DWT tramite la famiglia di funzioni $\psi_k^j(t)_{k,j \in \mathbb{Z}}$ che equivale a calcolare i coefficienti di dettaglio al variare di j e k nell'insieme finito di possibili valori:

$$\psi_k^j(t) = \int_{-\infty}^{+\infty} f(t) 2^{j/2} \psi(2^j t - k) dk \quad (2.1.7)$$

Analogamente, la realizzazione con la famiglia di funzioni $\varphi_k^j(t)_{k,j \in \mathbb{Z}}$ equivale a calcolare i coefficienti di approssimazione:

$$\varphi_k^j(t) = \int_{-\infty}^{+\infty} f(t) 2^{j/2} \varphi(2^j t - k) dk \quad (2.1.8)$$

La famiglia delle funzioni di scalatura φ per l'estrazione del contenuto di approssimazione del segnale e la famiglia delle funzioni wavelet ψ per l'estrazione dei dettagli del segnale sono state definite da Mallat stesso. Un approfondimento degli aspetti matematici e le dimostrazioni dei teoremi, solamente enunciati nel seguito, possono essere trovati in [16].

Tramite l'analisi multirisoluzione è possibile definire una successione di funzioni approssimanti convergenti alla funzione stessa. Le funzioni approssimanti saranno una combinazione lineare di funzioni ottenute da un insieme di funzioni di scaling derivanti da un'unica funzione di volta in volta traslata.

Definizione 1 Una analisi multirisoluzione di $L^2(\mathbb{R}^n)$ è una successione crescente di sottospazi lineari chiusi V_j , $j \in \mathbb{Z}$ di $L^2(\mathbb{R}^n)$:

$$\dots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \dots$$

che soddisfa le seguenti proprietà:

(i) l'intersezione è banale e l'unione è densa in $L^2(\mathbb{R}^n)$:

$$\bigcap_{j \in \mathbb{Z}} V_j = \{0\}, \quad \overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R}^n) \quad (2.1.9)$$

(ii) per ogni $f \in \mathbf{L}^2(\mathbb{R}^n)$ ed ogni $j \in \mathbb{Z}$,

$$f(x) \in V_j \leftrightarrow f(2x) \in V_{j+1} \quad (2.1.10)$$

(iii) per ogni $f \in \mathbf{L}^2(\mathbb{R}^n)$ ed ogni $k \in \mathbb{Z}^n$,

$$f(x) \in V_0 \leftrightarrow f(x - k) \in V_0 \quad (2.1.11)$$

(iv) esiste una funzione $g(x) \in V_0$ tale che $g(x - k)_{k \in \mathbb{Z}^n}$ è una base ortonormale di V_0 .

Enunciamo alcune proprietà dell'analisi multifunzione: dalla (ii) segue immediatamente che $f(x) \in V_0 \leftrightarrow f(2^j x) \in V_j$. Da (ii),(iii) e (iv) segue che $(2^{\frac{j}{2}} g(2^j x - k))_{k \in \mathbb{Z}^n}$ è una base ortonormale di V_j .

Definizione 2 Sia $\varphi(x) = g(x)$ e $\varphi_{j,k}(x) = 2^{\frac{j}{2}} \varphi(2^j x - k)$, φ è detta funzione di scaling dell'analisi di multirisoluzione.

È possibile rilassare la condizione (iv) richiedendo che $g(x - k)_{k \in \mathbb{Z}^n}$ sia una base di Riesz.

Si definisce una *base di Riesz* una famiglia di elementi le cui combinazioni lineari finite sono dense in V_0 ed esistono due costanti $C' > C > 0$ tali che per ogni successione di scalari $\alpha_0, \alpha_1, \dots$ si ha

$$C \left(\sum |\alpha_k|^2 \right)^{1/2} \leq \left\| \sum \alpha_k e_k \right\| \leq C' \left(\sum |\alpha_k|^2 \right)^{1/2}.$$

Definizione 3 In questo caso la costruzione di φ è più complessa e per i dettagli si rimanda a [17]

Sia $f \in \mathbf{L}^2(\mathbb{R}^n)$. L'analisi multirisoluzione permette di vedere f a diversi gradi di approssimazione: poiché i sottospazi $\{V_j\}$ sono chiusi, è possibile considerare la proiezione ortogonale di f su V_j per diversi j .

Definizione 4 Sia P_j l'operatore di proiezione ortogonale su V_j . $P_j f$ è detta approssimazione di f a livello j .

Le caratteristiche degli spazi V_j sono tali che è possibile costruire approssimatori di funzione, definendo opportune proiezioni su questi spazi. Poiché $\cup V_j$ è densa in $\mathbf{L}^2(\mathbb{R}^n)$ ogni funzione $f(x) \in \mathbf{L}^2(\mathbb{R}^n)$ può essere approssimata dalle proiezioni $P_j |f(x)|$ ovvero:

$$\forall f(x) \in \mathbf{L}^2(\mathbb{R}^n), \quad \lim_{j \rightarrow \infty} P_j [f(x)] = f(x) \quad (2.1.12)$$

Poiché $P_j[f(x)] = f(x)$ è un elemento di V_j , esso può essere scritto come combinazione lineare delle basi di V_j :

$$P_j[f(x)] = f(x) = \sum_k \lambda_{j,k} \varphi_{j,k}(x) \quad (2.1.13)$$

con appropriati coefficienti $\lambda_{j,k}$.

La differenza $P_{j+1}f - P_jf$ consiste nel dettaglio perso proiettando da V_{j+1} a V_j . La proprietà fondamentale della analisi multirisoluzione è che è possibile rappresentare questa differenza in una base di wavelet ortonormali.

Base wavelet e spazi di dettaglio

Definendo W_j come il complemento ortogonale di V_j in V_{j+1} :

$$V_{j+1} = V_j \oplus W_j = \nu_j + w_j \quad \text{con} \quad \nu_j \in V_j, w_j \in W_j$$

segue che per $j < J$

$$V_J = V_j \oplus_{i=j+1}^J W_i \quad (2.1.14)$$

Allora per costruzione V_{j+1} contiene tutte le informazioni di V_j . In altri termini, l'approssimazione della funzione in V_{j+1} avrà un livello di dettaglio maggiore rispetto all'approssimazione $P_j|f(x)|$. Di conseguenza, essendo W_j il complemento ortogonale di V_j , conterrà i dettagli per passare da un'approssimazione a risoluzione j a risoluzione maggiore.

Una wavelet è una funzione $\psi(x)$ per descrivere il dettaglio, se l'insieme delle funzioni $\psi(x - k)$ con $k \in \mathbb{Z}$ è una base di Riesz per W_0 . Omettendone la dimostrazione è possibile affermare che l'insieme delle funzioni $\psi_{j,k}(x)$ con $j, k \in \mathbb{Z}$ tale che :

$$\psi_{j,k}(x) = \sqrt{2^j} \psi(2^j x - k) \quad (2.1.15)$$

è una base di Riesz per $\mathbf{L}^2(\mathbb{R})$. Inoltre $\psi(x)$ sarà un elemento di V_1 dato che

$$V_1 = V_0 \oplus W_0 \quad \text{e} \quad W_0 \subseteq V_1.$$

Definiamo *wavelet madre* la funzione $\psi_{0,0}(x)$ per cui esiste una sequenza $g_k \in \mathbf{L}^2(\mathbb{R})$, tale che:

$$\psi_{0,0}(x) = \psi(x) = 2 \sum_k g_k \varphi(2x - k)$$

Dati i complementari V_j^c e W_j^c di V_j e W_j si ha che $V_j^c \bigoplus_{i=j}^{\infty} W_0$ e $W_j^c \bigoplus_{i \neq j}^{\infty} W_i$.

Definendo l'operatore Q_j su W_j parallelo a W_j^c , possiamo scrivere $f(x) \in \mathbf{L}^2(\mathbb{R}^n)$ nel seguente modo:

$$f(x) = \sum_j Q_j[f(x)] = \sum_{j,k} \gamma_{j,k} \psi_{j,k}(x)$$

Quest'ultima formulazione di $f(x)$ esprime la possibilità di ottenere il segnale come combinazione lineare delle funzioni dettaglio, una volta ottenuti i coefficienti $\gamma_{j,k}$.

Dalla scelta di una base ortogonale per la rappresentazione della funzione e dalle proprietà della multirisoluzione sarà possibile calcolare i coefficienti $\lambda_{j,k}$ come proiezione della funzione $f(x)$ e sulle corrispondenti funzioni di scaling $\lambda_{j,k} = \langle f, \varphi_{j,k} \rangle$, permettendoci di riscrivere la (2.1.13) come:

$$P_j[f(x)] = \sum_k \langle f, \varphi_{j,k} \rangle \varphi_{j,k}(x)$$

Analogamente le proiezioni Q_j sono ortogonali qualora le funzioni wavelet da cui derivano sono una base ortogonale di $\mathbf{L}^2(\mathbb{R}^n)$. Sarà possibile calcolare i coefficienti $\gamma_{j,k} = \langle f, \psi_{j,k} \rangle$ e esprimere i dettagli della scala j come:

$$Q_j[f(x)] = \sum_k \langle f, \psi_{j,k} \rangle \psi_{j,k}(x) = \sum_k \gamma_{j,k} \psi_{j,k}(x)$$

Inoltre P_j e Q_j permettono di ottenere le migliori approssimazioni di $f(x)$ nei rispettivi spazi.

$$\begin{aligned} f(x) &= \sum_j Q_j[f(x)] = \sum_{j,k} \langle f, \psi_{j,k} \rangle \psi_{j,k} = \\ \lim_{j \rightarrow \infty} P_j[f(x)] &= \lim_{j \rightarrow \infty} \sum_k \langle f, \varphi_{j,k} \rangle \varphi_{j,k}(x) \end{aligned}$$

Considerando wavelet biortogonali, piuttosto che ortogonali, è possibile ottenere coppie di filtri che permettono di calcolare la trasformata. L'esistenza di funzioni duali per le funzioni di scalatura e wavelet, $\tilde{\varphi}(x)$ e $\tilde{\psi}(x)$, permette di generare una multirisoluzione con sottospazi tali che:

$$\tilde{V}_j \perp W_j \quad \text{e} \quad V_j \perp \tilde{W}_j \tag{2.1.16}$$

Visto che $V_J = V_{j-1} \bigoplus W_{j-1}$, avremo:

$$\tilde{W}_j \perp (V_{j-i} \bigoplus W_{j-1})$$

e di conseguenza:

$$W_j \perp \tilde{W}_k \quad \text{per} \quad j \neq k$$

Il vantaggio delle funzioni duali consiste nella loro intercambiabilità con le primali, le cui definizioni sono simili. In questo modo la multirisoluzione è detta duale ed è espressa come segue:

$$P_j[f(x)] = \sum_k \langle f, \tilde{\varphi}_{j,k} \rangle \varphi_{j,k}(x)$$

$$Q_j[f(x)] = \sum_k \langle f, \tilde{\psi}_{j,k} \rangle \psi_{j,k}(x)$$

e la trasformata wavelet discreta potrà essere espressa nella forma:

$$f = \sum_{j,k} \langle f, \tilde{\psi}_{j,k} \rangle \psi_{j,k}$$

Per la dualità della multirisoluzione, è possibile analogamente esprimere la coppia P_j e Q_j nelle loro forme duali $\tilde{P}_j[f(x)]$ e $\tilde{Q}_j[f(x)]$.

Date le relazioni che legano gli spazi $\tilde{V}_j \perp W_j$ e $V_j \perp \tilde{W}_j$, vale la proprietà:

$$\langle \tilde{\varphi}(x), \psi(x-k) \rangle = \langle \tilde{\psi}, \varphi(x-k) \rangle = 0$$

Imponendo le condizioni:

$$\langle \tilde{\varphi}_{j,k}, \varphi_{j,k} \rangle = \delta_{l-k} \text{ e } \langle \tilde{\psi}_{j,l}, \psi_{i,k} \rangle = \delta_{j-i} \delta_{l-k} \text{ con } i, j, k, l \in \mathbb{Z}$$

in particolare

$$\langle \tilde{\varphi}(x), \varphi(-k) \rangle = \delta_k \text{ e } \langle \tilde{\psi}(x), \psi(-k) \rangle = \delta_k$$

2.1.5 Trasformata wavelet veloce

L'algoritmo della trasformata wavelet veloce elabora il segnale sfruttando le relazioni tra le funzioni base, precedentemente descritte, e riformula l'operazione di proiezione come l'applicazione di un filtro passa-basso e un filtro passa-alto. Possiamo descrivere una funzione $v(x) \in V_j$ nel seguente modo:

$$v_j(x) = \sum_k \lambda_{j,k} \varphi_{j,k}(x) = v_{j-1}(x) + w_{j-1}(x) =$$

$$= \sum_k \lambda_{j-1,k} \varphi_{j-1,k} + \sum_k \gamma_{j-1,k} \psi_{j-1,k}(x) \quad (2.1.17)$$

dove $v_{j-1}(x) \in V_{j-1}$ e $w_{j-1}(x) \in W_{j-1}$ con opportuni coefficienti $\lambda_{j,k}$, $\lambda_{j-1,k}$ e $\gamma_{j-1,k}$. Per le proprietà precedentemente descritte si avrà:

$$\lambda_{j-1,l} = \langle v_j, \tilde{\varphi}_{j-1,l} \rangle = \sqrt{2} \langle v_j, \sum_k \tilde{h}_{k-2l} \tilde{\varphi}_{j,l} \rangle = \sqrt{2} \sum_k \tilde{h}_{k-2l} \lambda_{j,l} \quad (2.1.18)$$

In modo analogo si ottiene:

$$\gamma_{j-1,l} = \sqrt{2} \sum_k \tilde{g}_{k-2l} \lambda_{j,l} \quad (2.1.19)$$

La trasformata wavelet veloce diretta è definita dall'applicazione ricorsiva di (2.1.18) e (2.1.19) tramite le quali è possibile calcolare i coefficienti di approssimazione e dettaglio a livello $(j-1)$. Tale procedimento equivale alla convoluzione dei vettori dei filtri passa-basso \tilde{h} e passa-alto \tilde{g} con il vettore dei coefficienti di approssimazione λ_j seguito da un sottocampionamento per eliminare la ridondanza.

La trasformata wavelet veloce inversa è ottenuta dall'applicazione ricorsiva della seguente:

$$\lambda_{j,k} = \sqrt{2} \sum_l h_{k-2l} \lambda_{j-1,l} + \sqrt{2} \sum_l h_{k-2l} g_{k-2l} \gamma_{j-1,l} \quad (2.1.20)$$

che consiste nella somma della convoluzione dei vettori h e g con le successioni ottenute da λ_{j-1} e γ_{j-1} .

Le due trasformate, diretta e inversa, definiscono *l'algoritmo di analisi e sintesi* [12].

La funzione di scalatura, detta anche funzione padre, $\varphi(x) \equiv \varphi_{0,0}(x) \in V_0 \subset V_1$ può essere riscritta come combinazione lineare secondo opportuni h_k delle funzioni base che generano V_1 :

$$\varphi(x) = \sum_k h_k \varphi_{1,k}(x) = \sqrt{2} \sum_k h_k \varphi(2x - k) \quad (2.1.21)$$

dove

$$h_k = \langle \varphi(x), \varphi_{1,k}(x) \rangle = \sqrt{2} \int_{-\infty}^{+\infty} \varphi(x) \varphi(2x - k) dx \quad (2.1.22)$$

In modo analogo per la funzione madre:

$$\psi(x) = \sum_k g_k \varphi_{1,k}(x) = \sqrt{2} \sum_k g_k \varphi(2x - k) \quad (2.1.23)$$

dove

$$g_k = \langle \psi(x), \varphi_{1,k}(x) \rangle = \sqrt{2} \int_{-\infty}^{+\infty} \varphi(x) \varphi(2x - k) dx \quad (2.1.24)$$

La (2.1.21) e (2.1.23) corrispondono rispettivamente alla prima e alla seconda equazione dell'analisi multirisoluzione. Inoltre è possibile ricavare la relazione tra i coefficienti $\lambda_{1,k}$ ed i

coefficienti $\lambda_{0,k}$ e $\gamma_{0,k}$.

$$\begin{aligned}\lambda_{0,k} &= \langle f(x), \varphi_{0,k}(x) \rangle = \langle \sum_l \lambda_{1,l} \varphi_{1,l}(x), \varphi_{0,k}(x) \rangle \\ &= \sum_l \lambda_{1,l} \langle \varphi_{1,k}(x), \varphi_{0,k}(x) \rangle\end{aligned}\quad (2.1.25)$$

ma $\langle \varphi_{1,k}(x), \varphi_{0,k}(x) \rangle = \int_{-\infty}^{+\infty} \sqrt{2} \varphi(2x-k) \varphi(x) dx$ e per sostituzione di variabili $x \rightarrow x+k$, confrontandola con (2.1.22) si ottiene:

$$\langle \varphi_{1,k}(x), \varphi_{0,k}(x) \rangle = \int_{-\infty}^{+\infty} \sqrt{2} \varphi(2x - (l-2k)) \varphi(x) dx = h_{l-2k}.\quad (2.1.26)$$

In conclusione:

$$\lambda_{0,k} = \sum_l \lambda_{1,l} h_{l-2k} = \sum_l \lambda_{1,l} \tilde{h}_{2k-l}\quad (2.1.27)$$

Pertanto la sequenza di coefficienti $\lambda_{0,k}$ si ottiene mediante filtraggio e decimazione per 2 della sequenza $\lambda_{1,k}$. Utilizzando (2.1.23) e (2.1.22) e con calcoli simili ai precedenti otteniamo:

$$\gamma_{0,n} = \sum_k \lambda_{1,k} g_{k-2n} = \sum_k \lambda_{1,k} \tilde{g}_{2n-k}\quad (2.1.28)$$

Iterando il procedimento si verifica la relazione tra i coefficienti $\lambda_{2,k}$ e i coefficienti $\lambda_{1,k}$ e $\gamma_{1,k}$:

$$\begin{aligned}\lambda_{1,k} &= \langle f(x), \varphi_{1,k}(x) \rangle = \langle \sum_l \lambda_{2,l} \varphi_{2,l}(x), \varphi_{1,k}(x) \rangle \\ &= \sum_l \lambda_{2,l} \langle \varphi_{2,k}(x), \varphi_{1,k}(x) \rangle\end{aligned}\quad (2.1.29)$$

dove $\langle \varphi_{2,k}(x), \varphi_{1,k}(x) \rangle = \int_{-\infty}^{+\infty} 2\varphi(4x-k) \sqrt{2} \varphi(2x-k) dx$. Per sostituzione di variabili $x \rightarrow x+k/2$ e confronto con (2.1.22) si ha:

$$\langle \varphi_{2,k}(x), \varphi_{1,k}(x) \rangle = \int_{-\infty}^{+\infty} \sqrt{2} \varphi(2x - (l-2k)) \varphi(x) dx = h_{l-2k}.\quad (2.1.30)$$

Si osserva inoltre che modificando la scala non varia il legame tra i coefficienti:

$$\lambda_{0,k} = \sum_l \lambda_{1,l} h_{l-2k} = \sum_l \lambda_{1,l} \tilde{h}_{2k-l}\quad (2.1.31)$$

Si ottiene allora lo schema di decomposizione a due livelli riportato in Figura 2.8. Lo schema sopra esposto può essere iterato fino al numero di livelli di decomposizione desiderati.

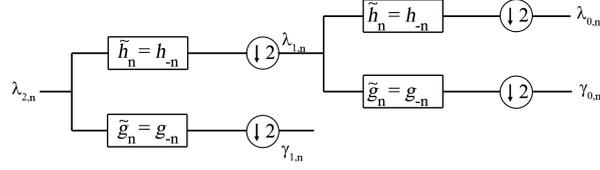


Figura 2.8: Decomposizione wavelet per 2 livelli.

Di seguito sono indicati i passi per la fase di ricostruzione, ricavando i coefficienti $\lambda_{1,k}$ a partire da $\lambda_{0,k}$ e $\gamma_{0,k}$ nell'ipotesi in cui $f(t) \in V_1$. Ricordando la (2.1.17), si ha:

$$\begin{aligned} \lambda_{1,k} &= \langle f(x), \varphi_{1,k}(t) \rangle = \langle \sum_l \lambda_{0,l} \varphi_{0,l}(x) + \sum_l \gamma_{0,l} \psi_{0,l}(x), \varphi_{1,k}(x) \rangle = \\ &= \sum_l \lambda_{0,l} \langle \varphi_{0,l}(x), \varphi_{1,k}(t) \rangle + \sum_l \gamma_{0,l} \langle \psi_{0,l}(x), \varphi_{1,k}(x) \rangle \end{aligned} \quad (2.1.32)$$

dove:

$$\langle \varphi_{0,l}(x), \varphi_{1,k}(x) \rangle = \int_{-\infty}^{+\infty} \varphi(x-l) \sqrt{2} \varphi(2x-k) dx = h_{k-2l}$$

e

$$\langle \psi_{0,l}(x), \varphi_{1,k}(x) \rangle = \int_{-\infty}^{+\infty} \psi(x-l) \sqrt{2} \varphi(2x-k) dx = g_{k-2l}$$

e infine avremo:

$$\lambda_{1,k} = \sum_l \lambda_{0,l} h_{k-2l} + \sum_l \gamma_{0,l} g_{k-2l} \quad (2.1.33)$$

Questa relazione mostra che è necessario realizzare un'espansione per 2 seguita da un filtraggio, i cui coefficienti sono proprio quelli determinati dalle due equazioni dell'analisi multirisoluzione (Figura 2.9). Lo schema può essere iterato così come già visto in fase di analisi all'aumentare del numero di livelli di decomposizione. Chiaramente è necessario imporre un vincolo alla lunghezza del segnale che deve essere una potenza di 2.

2.1.6 Caso bidimensionale

La trasformata wavelet può essere generalizzata al caso bidimensionale come nelle applicazioni che riguardano l'elaborazione di immagini. La trattazione già condotta per la wavelet monodimensionale si estende in modo analogo al caso bidimensionale e si dimostra che le

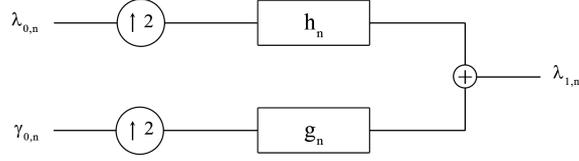


Figura 2.9: Implementazione di banco di filtri.

funzioni madre e padre sono separabili, per cui è possibile operare prima lungo le righe e poi lungo le colonne del segnale.

Sia $f(x, y) \in \mathbf{L}^2(\mathbb{R})$ un segnale bidimensionale: si definisce approssimazione multirisoluzione $(V_j)_{j \in \mathbb{Z}}$ di $\mathbf{L}^2(\mathbb{R})$ una sequenza di sottospazi di $\mathbf{L}^2(\mathbb{R})$ per la quale esiste una funzione $\varphi(x, y)$ che genera l'insieme delle funzioni $\{\varphi_{j,k,l} : j, k, l \in \mathbb{Z}\}$ definite come:

$$\varphi_{j,k,l}(x, y) = \sqrt{2^j} \varphi(2^j x - k, 2^j y - l)$$

i quali formano una base di Riesz per lo spazio V^j , soddisfacendo così i requisiti richiesti dal caso monodimensionale.

Si dice inoltre che V_j è separabile se $V_j = V_j^1 \otimes V_j^1$, ovvero se è decomponibile nel prodotto tensoriale di due sottospazi identici di $\mathbf{L}^2(\mathbb{R})$.

Di seguito sarà riportato il teorema che prova che una base di W_j può essere costruita scalando e traslando tre funzioni wavelet $\Psi^1(x, y) = \phi(x)\psi(y)$, $\Psi^2(x, y) = \psi(x)\phi(y)$ e $\Psi^3(x, y) = \psi(x)\psi(y)$.

Teorema 1 Siano $(V_{2^j})_{j \in \mathbb{Z}}$ una approssimazione multirisoluzione separabile di $\mathbf{L}^2(\mathbb{R})$, $\Phi(x, y) = \phi(x)\phi(y)$ la funzione di scala associata e $\psi(x)$ la wavelet mono-dimensionale associata alla funzione di scala $\phi(x)$.

Allora esistono tre wavelets:

$$\Psi^1(x, y) = \phi(x)\psi(y) \tag{2.1.34}$$

$$\Psi^2(x, y) = \psi(x)\phi(y) \tag{2.1.35}$$

$$\Psi^3(x, y) = \psi(x)\psi(y) \tag{2.1.36}$$

tali che, una base ortonormale di W_j è data da:

$$\begin{aligned} & (2^{-j}\Psi_j^1(x - 2^{-j}n, y - 2^{-j}m), \\ & 2^{-j}\Psi_j^2(x - 2^{-j}n, y - 2^{-j}m), \\ & 2^{-j}\Psi_j^3(x - 2^{-j}n, y - 2^{-j}m))_{(n,m)\in\mathbb{Z}^2} \end{aligned} \quad (2.1.37)$$

Come per il caso mono-dimensionale, la differenza tra $P_{j+1}f$ e P_jf è data dalla proiezione ortogonale del segnale su W_{2^j} ed è caratterizzata dai prodotti interni del segnale con la base ortonormale di W_{2^j} .

$$Q_j^1f = (\langle f(x, y), \Psi_2^1(x - 2^{-j}n, y - 2^{-j}m) \rangle)_{(n,m)\in\mathbb{Z}^2} \quad (2.1.38)$$

$$Q_j^2f = (\langle f(x, y), \Psi_2^2(x - 2^{-j}n, y - 2^{-j}m) \rangle)_{(n,m)\in\mathbb{Z}^2} \quad (2.1.39)$$

$$Q_j^3f = (\langle f(x, y), \Psi_2^3(x - 2^{-j}n, y - 2^{-j}m) \rangle)_{(n,m)\in\mathbb{Z}^2} \quad (2.1.40)$$

In particolare $Q_{2^j}^1f$ rappresenta il dettaglio discreto orizzontale alla risoluzione j , perché calcolato come prodotto interno del segnale $f(x, y)$ e la wavelet $\Psi^1(x, y)$. In modo analogo $Q_{2^j}^2f$ e $Q_{2^j}^3f$ rappresentano il dettaglio discreto orizzontale e diagonale rispettivamente.

Come per il caso mono dimensionale il calcolo delle approssimazioni discrete saranno riconducibili ad operazione di filtraggio e campionamento.

$$P_jf = ((f(x, y) * \phi_{2^j}(-x)\phi_{2^j}(-y))(2^{-j}n, 2^{-j}m))_{(n,m)\in\mathbb{Z}^2} \quad (2.1.41)$$

$$Q_j^1f = ((f(x, y) * \phi_{2^j}(-x)\psi_{2^j}(-y))(2^{-j}n, 2^{-j}m))_{(n,m)\in\mathbb{Z}^2} \quad (2.1.42)$$

$$Q_j^2f = ((f(x, y) * \psi_{2^j}(-x)\phi_{2^j}(-y))(2^{-j}n, 2^{-j}m))_{(n,m)\in\mathbb{Z}^2} \quad (2.1.43)$$

$$Q_j^3f = ((f(x, y) * \psi_{2^j}(-x)\psi_{2^j}(-y))(2^{-j}n, 2^{-j}m))_{(n,m)\in\mathbb{Z}^2} \quad (2.1.44)$$

Nel caso multidimensionale ci si riferisce generalmente alla DWT con il nome di decomposizione wavelet, in cui righe e colonne vengono considerate separatamente (Figura 2.10). Esistono due tipologie principali di decomposizione:

- Decomposizione standard: si considerano prima le righe dell'immagine, come sequenze distinte e su ciascuna di essa viene applicata una DWT monodimensionale fino al massimo livello, ovvero fino a quando non è più possibile effettuare una decimazione dei campioni. Lo stesso procedimento è applicato successivamente alle colonne (Figura 2.10-*sinistra*).

- Decomposizione non-standard: anche in questo caso vengono considerate separatamente righe e colonne, ma si procede alternando la trasformata wavelet sulle righe e sulle colonne ad ogni passo (Figura 2.10-*destra*).

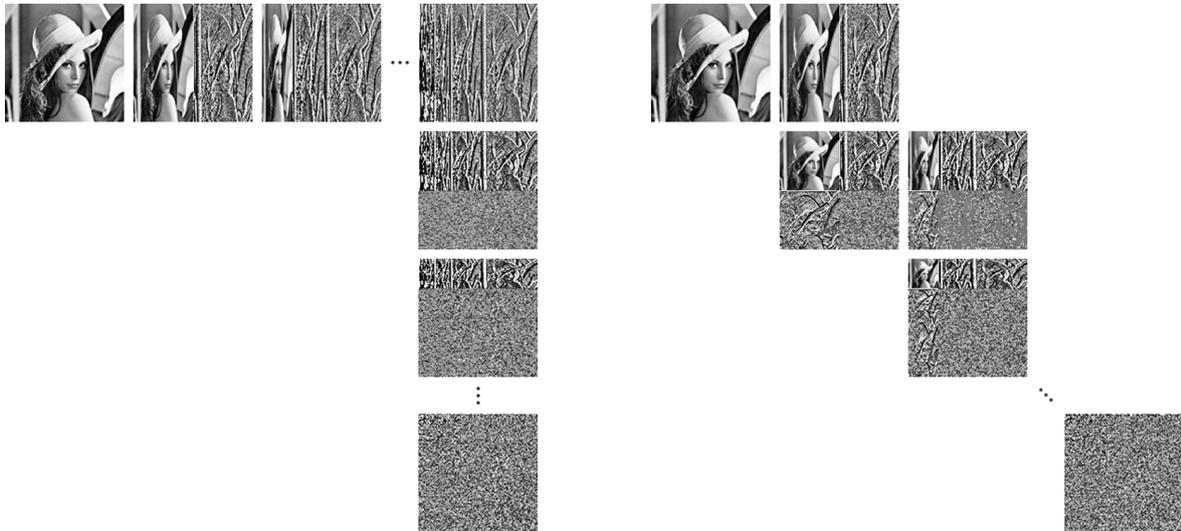


Figura 2.10: Esempio di decomposizione standard (sinistra) e non-standard (destra).

2.1.7 La trasformata *à trous*

Una limitazione della trasformata finora descritta è l'invarianza per traslazione del segnale. Tale limitazione è dovuta al sottocampionamento che viene eseguito in fase di analisi, in cui i coefficienti sono calcolati a partire da un sottoinsieme di pixel dell'immagine, perdendo talvolta alcune informazioni legate alla localizzazione spaziale delle componenti di frequenza. Sono state sviluppate metodiche, note come Undecimated Wavelet Transform (UWT) in cui tale limitazione è stata superata a discapito della complessità computazionale [18, 19].

All'immagine in ingresso viene applicata una versione modificata della trasformata wavelet per cui il segnale non viene sottoposto a sottocampionamento lungo la decomposizione e durante l'intero processo vengono mantenute molte informazioni ridondanti. Lo schema si applica tramite opportuni filtri passa-basso e passa-alto, in modo tale da non effettuare il sottocampionamento. In questo caso vengono inseriti tra ogni coppia di coefficienti un numero di zeri proporzionale al livello (questi valori nulli rappresentano i "buchi" che danno il nome all'algorithm) (Figura 2.11).

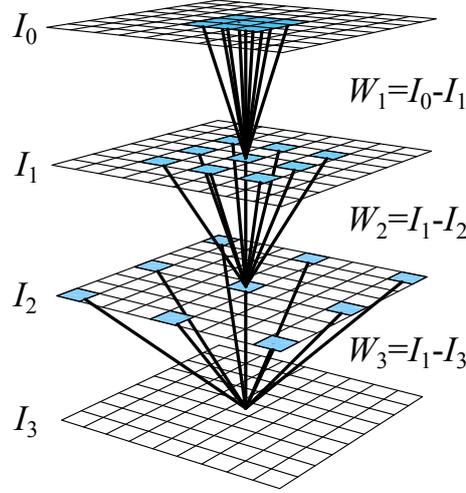


Figura 2.11: Kernel con buchi.

L'introduzione di questi elementi nulli simula il restringimento dell'immagine; più che una struttura piramidale, può essere vista come una struttura a forma di parallelepipedo in cui gli unici elementi che si propagano a piramide sono gli elementi non nulli del kernel.

Assumendo che l'immagine di input I sia ottenibile come prodotto scalare:

$$I_0(\mathbf{p}) = I(\mathbf{p}) = \langle f(\mathbf{q}), \phi(\mathbf{q} - \mathbf{p}) \rangle \quad (2.1.45)$$

dove ϕ indica la funzione di scalatura. Nel caso diadico:

$$\frac{1}{4} \phi\left(\frac{\mathbf{p}}{2^i}\right) = \sum_{\mathbf{q} \in \mathbb{D}_\infty} \ell_{\mathbf{q}} \phi\left(\frac{\mathbf{p}}{2^{i-1}} - \mathbf{q}\right) \quad (2.1.46)$$

dove ℓ è un filtro passa-basso. Le approssimazioni successive I_i di I_0 possono essere calcolate da:

$$I_i(\mathbf{p}) = \frac{1}{4^i} \left\langle f(\mathbf{q}), \phi\left(\frac{\mathbf{q} - \mathbf{p}}{2^i}\right) \right\rangle = \sum_{\mathbf{q} \in \mathbb{D}_\infty} \ell_{\mathbf{q}} I_{i-1}(\mathbf{p} + 2^{i-1} \mathbf{q}). \quad (2.1.47)$$

I coefficienti wavelet W_i sono ottenuti dalla convoluzione di I_i con un filtro passa-alto h :

$$W_i(\mathbf{p}) = \frac{1}{4^i} \left\langle f(\mathbf{q}), \psi\left(\frac{\mathbf{q} - \mathbf{p}}{2^i}\right) \right\rangle = \sum_{\mathbf{q} \in \mathbb{D}_\infty} h_{\mathbf{q}} I_{i-1}(\mathbf{p} + 2^{i-1} \mathbf{q}) \quad (2.1.48)$$

dove ψ è la funzione wavelet definita come:

$$\frac{1}{4^i} \psi\left(\frac{\mathbf{p}}{2^i}\right) = \frac{1}{4^{i-1}} \sum_{\mathbf{q} \in \mathbb{D}_\infty} h_{\mathbf{q}} \phi\left(\frac{\mathbf{p}}{2^{i-1}} - \mathbf{q}\right).$$

Il filtro passa-alto più semplice è dato dalla differenza tra due approssimazioni di scala consecutive:

$$W_i(\mathbf{p}) = I_{i-1}(\mathbf{p}) - I_i(\mathbf{p}) \quad (2.1.49)$$

il quale può essere riscritto, confrontando le equazioni (2.1.47) e (2.1.48) come:

$$\frac{1}{4}\psi\left(\frac{\mathbf{p}}{2}\right) = \phi(\mathbf{p}) - \frac{1}{4}\phi\left(\frac{\mathbf{p}}{2}\right).$$

Imponiamo la B_s -spline $\phi = B_s$:

$$B_s(z) = \frac{1}{2s!} \sum_{t=0}^{s+1} (-1)^t \binom{t}{s+1} \left| z + t - \frac{s+1}{2} \right|^s.$$

e nel caso specifico dell'interpolazione bilineare in cui $s = 1$, (Figg. 2.12, 2.13) per la separabilità delle variabili ($\phi(\mathbf{p}) = \phi(x)\phi(y)$) si ottiene:

$$\phi(\mathbf{p}) = \frac{1}{4}(|x-1| - 2|x| + |x+1|)(|y-1| - 2|y| + |y+1|).$$

Limitando D_∞ ad un disco di raggio $\sqrt{2}$ e confrontando il polinomio precedente e la funzione di scalatura (2.1.46) si avrà per $i = 0$:

$$\phi(\mathbf{p}) = \sum_{\mathbf{q} \in \mathbb{D}_{\sqrt{2}}} 4\ell_{\mathbf{q}} \phi(2\mathbf{p} - \mathbf{q})$$

ottenendo un sistema di 25 equazioni in 9 incognite la cui soluzione:

$$\ell = \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}.$$

Sfruttando nuovamente la separabilità delle variabili $h(\mathbf{p}) = h(x)h(y)$, il kernel bidimensionale può essere semplificato alla sua versione mono dimensionale, da applicare lungo x e y :

$$\ell = \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \end{pmatrix}.$$

Ad ogni iterazione i della trasformata à trous la dimensione di ℓ non cambia, ma cambia la distanza tra il pixel centrale e quello adiacente e sarà uguale a 2^{i-1} .

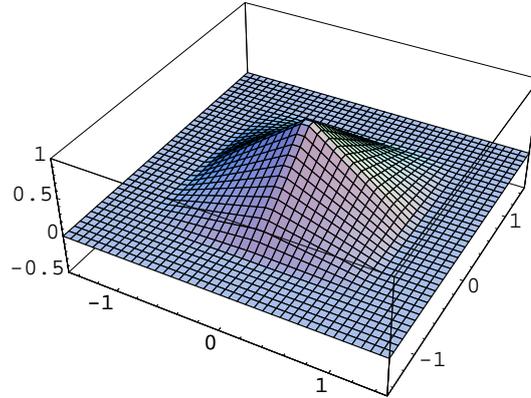


Figura 2.12: Funzione di scalatura corrispondente a B_1 .

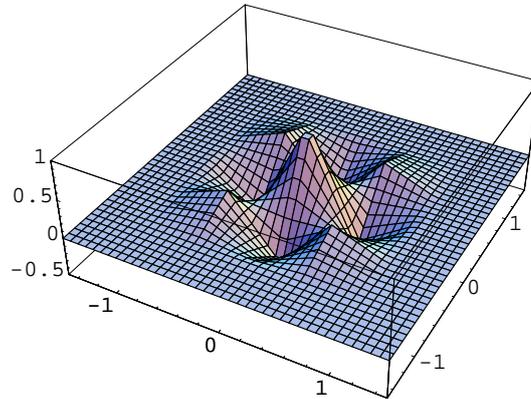


Figura 2.13: Funzione wavelet corrispondente a B_1 .

Di seguito i passi che determinano i coefficienti di :

$$\ell = \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}.$$

L'espansione della funzione di scalatura come interpolazione lineare:

$$\begin{aligned} \phi(\mathbf{p}) &= \frac{1}{4}(|x-1| - 2|x| + |x+1|)(|y-1| - 2|y| + |y+1|) = \\ &= \frac{1}{4}|x-1||y-1| - \frac{1}{2}|x-1||y| + \frac{1}{4}|x-1||y+1| - \frac{1}{2}|x||y-1| + |x||y| \\ &= -\frac{1}{2}|x||y+1| + \frac{1}{4}|x+1||y-1| - \frac{1}{2}|x+1||y| + \frac{1}{4}|x+1||y+1|. \end{aligned}$$

è valutata su una finestra $D_{\sqrt{2}} 3 \times 3$:

$$\begin{aligned}
\phi(\mathbf{p}) = & \sum_{x_q, y_q = -1}^1 4\ell_{\mathbf{q}} \phi(2\mathbf{p} - \mathbf{q}) = |2x - 1| |2y - 1| (\ell_{00} - 2\ell_{01} - 2\ell_{10} \\
& + 4\ell_{11}) + 2|2x - 1| |y - 1| (\ell_{01} - 2\ell_{11}) + 2|2x - 1| |y| (\ell_{0-1} - 2\ell_{00} + \\
& \ell_{01} - 2\ell_{1-1} + 4\ell_{10} - 2\ell_{11}) + 2|2x - 1| |y + 1| (\ell_{0-1} - 2\ell_{1-1}) + |2x - 1| \\
& |2y + 1| (-2\ell_{0-1} + \ell_{00} + 4\ell_{1-1} - 2\ell_{10}) + 2|x - 1| |2y - 1| (\ell_{10} - 2\ell_{11}) \\
& + 4|x - 1| |y - 1| \ell_{11} + 4|x - 1| |y| (\ell_{1-1} - 2\ell_{10} + \ell_{11}) + 4|x - 1| |y + 1| \\
& \ell_{1-1} + 2|x - 1| |2y + 1| (-2\ell_{1-1} + \ell_{10}) + 2|x| |2y - 1| (\ell_{-10} - 2\ell_{-11} - \\
& 2\ell_{00} + 4\ell_{01} + \ell_{10} - 2\ell_{11}) + 4|x| |y - 1| (\ell_{-11} - 2\ell_{01} + \ell_{11}) + 4|x| |y| \\
& (\ell_{-1-1} - 2\ell_{-10} + \ell_{-11} - 2\ell_{0-1} + 4\ell_{00} - 2\ell_{01} + \ell_{1-1} - 2\ell_{10} + \ell_{11}) + \\
& 4|x| |y + 1| (\ell_{-1-1} - 2\ell_{0-1} + \ell_{1-1}) + 2|x| |2y + 1| (-2\ell_{-1-1} + \ell_{-10} + \\
& 4\ell_{0-1} - 2\ell_{00} - 2\ell_{1-1} + \ell_{10}) + 2|x + 1| |2y - 1| (\ell_{-10} - 2\ell_{-11}) + \\
& 4|x + 1| |y - 1| \ell_{-11} + 4|x + 1| |y| (\ell_{-1-1} - 2\ell_{-10} + \ell_{-11}) + 4|x + 1| \\
& |y + 1| \ell_{-1-1} + 2|x + 1| |2y + 1| (-2\ell_{-1-1} + \ell_{-10}) + |2x + 1| |2y - 1| \\
& (-2\ell_{-10} + 4\ell_{-11} + \ell_{00} - 2\ell_{01}) + 2|2x + 1| |y - 1| (-2\ell_{-11} + \ell_{01}) + \\
& 2|2x + 1| |y| (-2\ell_{-1-1} + 4\ell_{-10} - 2\ell_{-11} + \ell_{0-1} - 2\ell_{00} + \ell_{01}) + 2 \\
& |2x + 1| |y + 1| (-2\ell_{-1-1} + \ell_{0-1}) + |2x + 1| |2y + 1| (4\ell_{-1-1} - 2\ell_{-10} \\
& - 2\ell_{0-1} + \ell_{00}).
\end{aligned}$$

Il confronto dei due polinomi, valido per ogni pixel \mathbf{p} determina il seguente sistema:

$$\left\{ \begin{array}{l}
l_{00} - 2l_{01} - 2l_{10} + 4l_{11} = 0 \\
l_{01} - 2l_{11} = 0 \\
l_{0-1} - 2l_{00} + l_{01} - 2l_{1-1} + 4l_{10} - 2l_{11} = 0 \\
l_{0-1} - 2l_{1-1} = 0 \\
2l_{0-1} - l_{00} - 4l_{1-1} + 2l_{10} = 0 \\
l_{10} - 2l_{11} = 0 \\
16l_{11} - 1 = 0 \\
8(l_{1-1} - 2l_{10} + l_{11}) + 1 = 0 \\
16l_{1-1} - 1 = 0 \\
2l_{1-1} - l_{10} = 0 \\
l_{-10} - 2l_{-11} - 2l_{00} + 4l_{01} + l_{10} - 2l_{11} = 0 \\
8(l_{-11} - 2l_{01} + l_{11}) + 1 = 0 \\
4(l_{-1-1} - 2l_{-10} + l_{-11} - 2l_{0-1} + 4l_{00} - 2l_{01} + l_{1-1} - 2l_{10} + l_{11}) - 1 = 0 \\
8(l_{-1-1} - 2l_{0-1} + l_{1-1}) + 1 = 0 \\
2l_{-1-1} - l_{-10} - 4l_{0-1} + 2l_{00} + 2l_{1-1} - l_{10} = 0 \\
l_{-10} - 2l_{-11} = 0 \\
16l_{-11} - 1 = 0 \\
8(l_{-1-1} - 2l_{-10} + l_{-11}) + 1 = 0 \\
16l_{-1-1} - 1 = 0 \\
2l_{-1-1} - l_{-10} = 0 \\
2l_{-10} - 4l_{-11} - l_{00} + 2l_{01} = 0 \\
2l_{-11} - l_{01} = 0 \\
2l_{-1-1} - 4l_{-10} + 2l_{-11} - l_{0-1} + 2l_{00} - l_{01} = 0 \\
2l_{-1-1} - l_{0-1} = 0 \\
4l_{-1-1} - 2l_{-10} - 2l_{0-1} + l_{00} = 0
\end{array} \right.$$

il quale ammette le seguenti soluzioni:

$$l_{-1-1} = l_{-11} = l_{1-1} = l_{11} = \frac{1}{16}, \quad l_{-10} = l_{0-1} = l_{01} = l_{10} = \frac{1}{8}, \quad l_{00} = \frac{1}{4}.$$

2.2 Reti Neurali Artificiali

Le reti neurali artificiali (ANN Artificial Neural Network) nascono per soddisfare l'esigenza di individuare un modello matematico per la simulazione del sistema nervoso umano, in particolare la sua struttura e le relative funzioni di base. In analogia al sistema nervoso umano, formato da strutture tra loro interconnesse dette neuroni, le reti neurali sono anch'esse formate da neuroni artificiali, organizzati in una struttura a rete. Un richiamo alle principali caratteristiche dei neuroni e del loro funzionamento, permette di comprendere meglio i principi delle reti neurali artificiali. Un neurone è formato da tre parti principali: i dendriti, che rappresentano un insieme di fibre in ingresso, il soma, il corpo cellulare e l'assone, che è la diramazione principale, il quale si ramifica in differenti terminazione che permettono di essere in contatto con altri neuroni. Queste cellule nervose interconnesse tra loro compongono la parte fondamentale del cervello umano. La comunicazione tra i vari neuroni avviene attraverso la trasmissione di impulsi nelle zone di contatto, tali zone sono dette sinapsi (Figura 2.14).

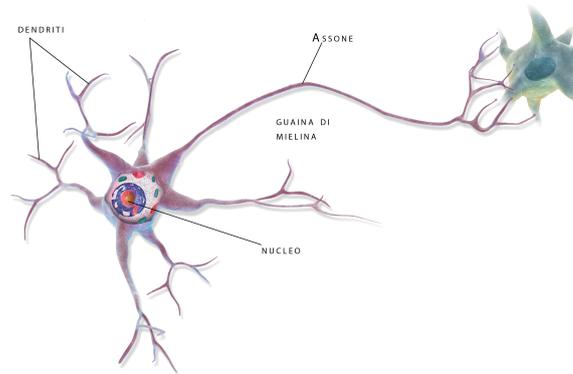


Figura 2.14: Neurone biologico.

Gli impulsi generati e trasmessi, fungono da inibitori o attivatori per i neuroni. I neuroni, organizzati in strutture, possono essere pensati come sottoinsiemi, ovvero moduli indipendenti in grado di svolgere specifici compiti; in questo modo più informazioni possono essere elaborate in parallelo e quindi più attività possono essere svolte contemporaneamente. Per analogia al sistema nervoso, nelle reti neurali artificiali il neurone artificiale sarà l'unità computazionale mentre le sinapsi saranno le interconnessioni, ed il modello e i processi che le governano possono essere visti come una semplificazione dei processi cerebrali. Il neuroscien-

ziato Warren McCulloch e il logico Walter Pitts nel 1943 proposero formalmente il primo modello di neurone artificiale nell'articolo: "A logical calculus of the ideas immanent in nervous activity" [20]. Il neurone artificiale (Figura 2.15) può essere descritto come un modello matematico, costituito da R canali di ingresso, ai quali sono associati dei pesi w_i , i quali assumono valori positivi o negativi a seconda della loro funzione eccitatoria o inibitoria.

I dati in input saranno espressi da un vettore \mathbf{p} mentre il vettore \mathbf{w} esprimerà il vettore dei pesi, entrambi avranno una dimensione pari a R :

$$\mathbf{p} = (p_1, p_2, p_3, \dots, p_R)^T, \quad \mathbf{w} = (w_1, w_2, w_3, \dots, w_R)^T$$

Computata nel corpo del neurone, la somma dei prodotti $p_i w_i$, viene confrontata con un valore di soglia t_0 cfr. (2.2.1). Il neurone restituirà in output un valore binario, pari a 1 se la somma supera il valore di soglia t_0 , 0 altrimenti.

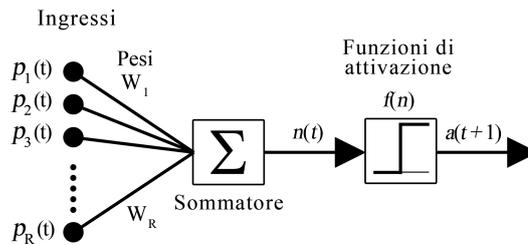


Figura 2.15: Neurone artificiale.

$$f(n) = \begin{cases} 1 & n \geq t_0 \\ 0 & n < t_0 \end{cases} \quad (2.2.1)$$

L'output del neurone è rappresentato dalla seguente espressione:

$$\mathbf{a} = f(n) = f\left(\sum_{i=0}^R w_i p_i\right)$$

Il modello proposto da McCulloch e Pitts, seppur innovativo per il periodo storico, fu sostituito dal modello proposto da Frank Rosenblatt nel 1958 [21], conosciuto come perceptrone (Figura 2.16).

Il modello descritto dal perceptrone, nella sua versione originale, è in grado di risolvere problemi di classificazione dicotomica. Analogamente al neurone artificiale, l'output sarà determinato dal valore restituito dalla funzione di attivazione, cfr. (2.2.2), nella quale il valore soglia t_0 , viene sostituito, introducendo un parametro di ingresso detto bias, il cui valore sarà

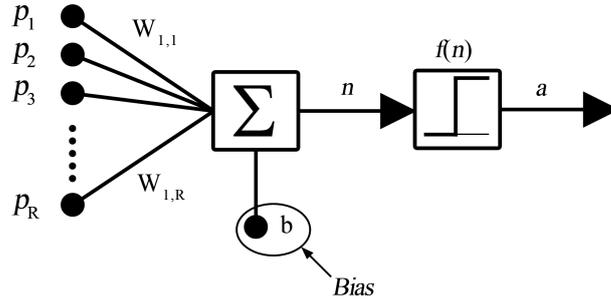


Figura 2.16: Il Perceptrone.

opposto a t_0 e sommato alla somma pesata degli input $\sum_{i=1}^R w_i p_i$, il quale varia secondo la propensione del neurone ad attivarsi, per variare la soglia di attivazione del neurone.

$$f(n) = \begin{cases} 1 & n \geq 0 \\ 0 & n < 0 \end{cases} \quad (2.2.2)$$

Considerando un ulteriore canale di input e con il relativo peso, definiti rispettivamente $p_0 = 1$ e $w_0 = b$, è quindi possibile esprimere in forma vettoriale l'output del perceptrone,

$$\mathbf{a} = f(n) = f\left(\sum_{i=1}^R (w_i p_i) + b_i\right) = f\left(\sum_{i=0}^R (w_i p_i)\right)$$

La funzione cfr. (2.2.2) è detta a gradino, tra le funzioni di attivazione è la più semplice, ma restituendo valori binari non rispecchia pienamente il modello biologico, inoltre non è una funzione derivabile; essa può essere sostituita con altre funzioni derivabili che, come vedremo successivamente permetteranno di risolvere alcuni problemi legati alla derivabilità della funzione di errore, quelle comunemente utilizzate sono mostrate in Figura 2.17:

Più perceptron possono cooperare, formando una rete di perceptron (Figura 2.18) in grado di svolgere compiti più complessi:

Trattandosi di più perceptron interconnessi, i pesi ad essi associati possono essere espressi in forma matriciale dalla matrice \mathbf{W} (2.2.3), di dimensioni $S \times R$, dove S esprime il numero di neuroni che compongono la rete. Il valore di output sarà espresso da un vettore di dimensione S (2.2.4), in cui ad ogni i -esimo elemento corrisponderà il valore di output dell' i -esimo neurone della rete, considerazioni analoghe valgono per l'elemento bias il quale sarà sostituito da un vettore di dimensione S (2.2.5).

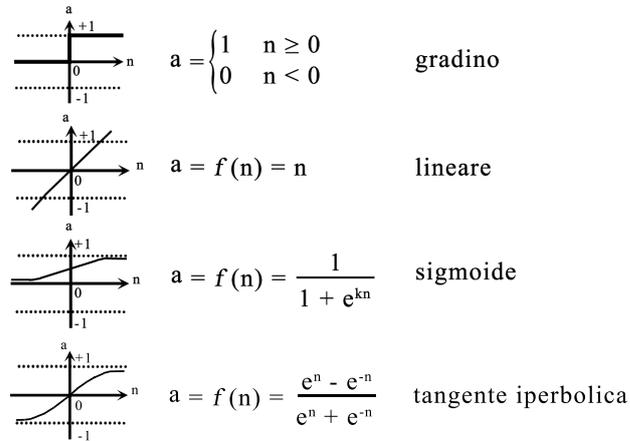


Figura 2.17: Alcune funzioni di attivazione.

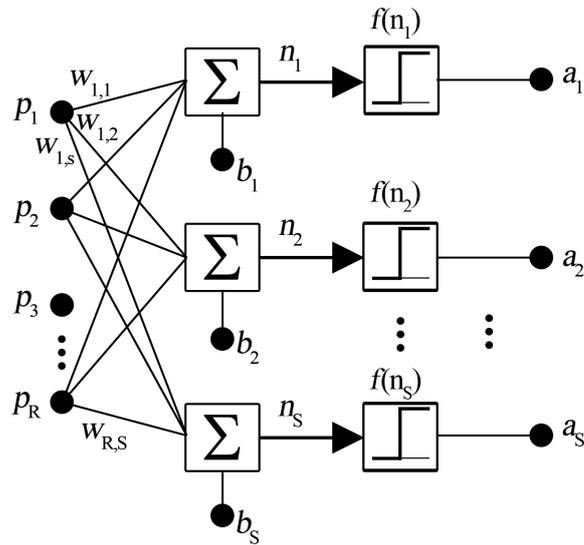


Figura 2.18: Rete di perceptroni.

$$\mathbf{W} = \begin{pmatrix} \mathbf{w}_{11} & \mathbf{w}_{12} & \mathbf{w}_{13} & \cdot & \mathbf{w}_{1R} \\ \mathbf{w}_{21} & \mathbf{w}_{22} & \mathbf{w}_{23} & \cdot & \mathbf{w}_{2R} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \mathbf{w}_{S1} & \mathbf{w}_{S2} & \mathbf{w}_{S3} & \cdot & \mathbf{w}_{SR} \end{pmatrix} \quad (2.2.3)$$

$$\mathbf{a} = (a_1, a_2, a_3, \cdot, \cdot, a_S)^T \quad \text{vettore di output} \quad (2.2.4)$$

$$\mathbf{b} = (b_1, b_2, b_3, \dots, b_S)^T \quad \text{vettore bias} \quad (2.2.5)$$

Analogamente al perceptrone, la funzione di attivazione sarà espressa dalla (2.2.2) e l'output della singola unità è espresso da:

$$\mathbf{a}_i = \mathbf{f}(\mathbf{n}_i) = \mathbf{f}(\mathbf{w}_i \mathbf{p} + b_i)$$

dove con \mathbf{w}_i indichiamo la i -esima riga della matrice dei pesi.

Si può già notare come il comportamento dell'output della rete possa essere condizionato dal valore dei pesi w e del bias. Peraltro, l'equazione $\mathbf{W}\mathbf{p} + \mathbf{b}$ definisce un iperpiano nello spazio euclideo a dimensione R . In questa considerazione, si assume che il vettore \mathbf{p} sia un vettore dello spazio euclideo a dimensione R , ovvero il caso in cui \mathbf{p} rappresenti un pattern descritto da attributi reali. Dal punto di vista geometrico, un iperpiano suddivide lo spazio in due semispazi e pertanto può essere utilizzato per effettuare una classificazione dicotomica.

La rete apprende alterando i pesi in modo tale da rafforzare le decisioni corrette e scoraggiare quelle non corrette. Si è in presenza di un apprendimento supervisionato se il valore di output per ogni elemento del training set è noto e l'aggiornamento dei pesi avviene confrontando l'output della rete con il valore atteso; altrimenti l'apprendimento si definisce non supervisionato e la rete apprenderà le similarità tra i dati forniti. La differenza tra il valore di output restituito dalla rete e il valore atteso (nel caso di apprendimento supervisionato) rappresenta la misura dell'errore \mathbf{E} , tale funzione è definita sulla matrice dei pesi, $\mathbf{E}(\mathbf{W})$, che variano nel tempo. Se \mathbf{E} è differenziabile è possibile utilizzare metodi numerici per trovare il valore \mathbf{W}^* che minimizza \mathbf{E} .

$$\mathbf{W}^* = \arg \min_{\mathbf{w}} \mathbf{E}(\mathbf{W}) \quad (2.2.6)$$

Uno degli algoritmi più conosciuti per il calcolo del minimo, nel caso in cui la funzione sia differenziabile è l'algoritmo del gradiente discendente [22]. Supponendo che \mathbf{E} sia l'errore quadratico totale avremo:

$$\mathbf{E}(\mathbf{W}) = \frac{1}{2} \sum_{q=1}^M \sum_{i=1}^S (\mathbf{a}_i^q - \mathbf{d}_i^q)^2 = \sum_{q=1}^M \mathbf{E}^q(\mathbf{W})$$

dove M indica il numero di input forniti alla rete. Mentre l'errore quadratico parziale commesso dalla rete sull'input q –esimo sarà:

$$E^q(\mathbf{W}) = \frac{1}{2} \sum_{i=1}^S (a_i^q - d_i^q)^2$$

dove con a^q e d^q indichiamo rispettivamente l'output della rete e il target (valore atteso) per il q –esimo elemento del training set (S il numero di neuroni). Il criterio di ricerca del minimo suggerito dall'algoritmo del gradiente discendente è quello di aggiornare i pesi nella direzione opposta al gradiente della funzione di errore.

$$\mathbf{W}' = \mathbf{W} - \eta \nabla E \mathbf{W}$$

dove \mathbf{W}' indica la matrice dei pesi aggiornata ed η il learning rate (definito in seguito). Dal punto di vista geometrico, $E^q(\mathbf{W})$ è definito come una funzione dipendente dai pesi w e rappresenta una funzione convessa, inoltre muoversi nella direzione opposta al gradiente significherà andare nella direzione del minimo (Figura 2.19):

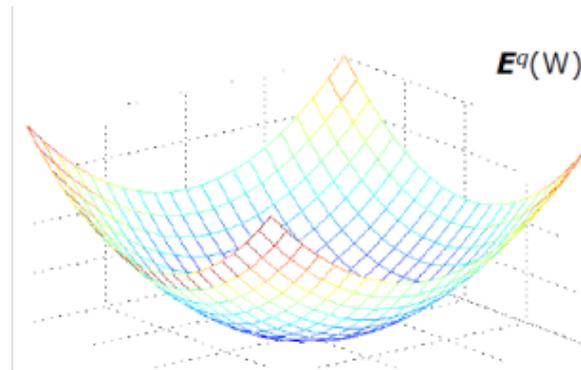


Figura 2.19: Rappresentazione geometrica della funzione di errore.

In letteratura i principali approcci per la valutazione dell'errore sono due: batch e incrementale. Nell'approccio batch i pesi vengono aggiornati solo quando saranno presentati alla rete tutti gli elementi che compongono il training set, ovvero alla fine di un epoca u , dove per epoca intendiamo un intero ciclo di apprendimento. La seguente espressione rappresenta l'aggiornamento dei pesi in un approccio batch:

$$w_{ik}(u+1) = w_{ik}(u) + \Delta w_{ik}(u)$$

$$\Delta w_{ik}(u) = -\eta \frac{\partial \mathbf{E}(u)}{\partial w_{ik}} = -\eta \sum_{q=1}^M \frac{\partial \mathbf{E}^q(u)}{\partial w_{ik}} = \sum_{q=1}^M \Delta w_{ik}^q(u)$$

ponendo

$$\Delta w_{ik}^q(u) = -\eta \frac{\partial \mathbf{E}^q(u)}{\partial w_{ik}}$$

Con η indichiamo il learning rate, il cui valore è compreso nell'intervallo $(0,1)$, esso determina quanto *grande* sarà l'aggiornamento dei pesi; la scelta del valore di learning non è del tutto banale, infatti grandi valori di η velocizzano l'apprendimento, rischiando però di non identificare il valore di minimo e quindi non far convergere l'algoritmo, di contro valori troppo piccoli di η potrebbero far convergere l'algoritmo troppo lentamente (Figura 2.20).

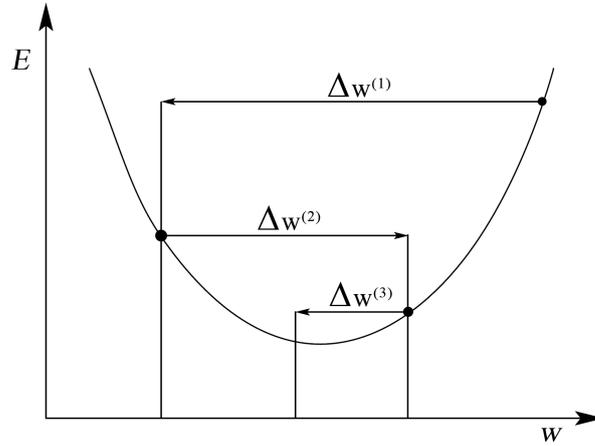


Figura 2.20: Learning rate.

Nell'approccio incrementale i pesi vengono aggiornati per ciascun elemento del training set presentato durante la singola epoca:

$$\Delta w_{ik}^{q+1}(u) = w_{ik}^q(u) + \Delta w_{ik}^q(u)$$

dove $\Delta w_{ik}^q(u) = -\eta \frac{\partial \mathbf{E}(u)^q}{\partial w_{ik}}$ e sostituendo \mathbf{a}_i con il prodotto $\sum_{k=0}^R w_{ik}^q(u) \mathbf{p}_k^q$ si ottiene:

$$\mathbf{E}^q(u) = \frac{1}{2} \sum_{i=1}^S (\mathbf{a}_i^q - \mathbf{d}_i^q)^2 = \frac{1}{2} \sum_{i=1}^S \left(\left(\sum_{k=0}^R w_{ik}^q(u) \mathbf{p}_k^q \right) - \mathbf{d}_i^q \right)^2$$

L'apprendimento incrementale (denominato anche *on-line*) comporta un apprendimento computazionalmente più lento, infatti l'aggiornamento dei pesi viene calcolato per ciascuno degli M elementi del training, ma risulta più accurato rispetto a quello batch.

Nel caso in cui la funzione di transizione f scelta sia differenziabile allora è possibile calcolare la derivata dell'errore.

$$\begin{aligned} \frac{\partial \mathbf{E}(u)}{\partial w_{ik}} &= \frac{\partial \frac{1}{2} \sum_{i=1}^S (\mathbf{a}_i^q - \mathbf{d}_i^q)^2}{\partial w_{ik}} = \frac{\partial \frac{1}{2} \sum_{i=1}^S \left(\left(\sum_{k=0}^R w_{ik}^q(u) \mathbf{p}_k^q \right) - \mathbf{d}_i^q \right)^2}{\partial w_{ik}} = \\ &= (\mathbf{a}_i^q - \mathbf{d}_i^q) \frac{\partial \left(\sum_{k=0}^R w_{ik}^q(u) \mathbf{p}_k^q \right)}{\partial w_{ik}} = (\mathbf{a}_i^q - \mathbf{d}_i^q) f'(n_i^q) = \delta_i^q p_k^q \end{aligned} \quad (2.2.7)$$

Ponendo $\delta_i^q p_k^q = (\mathbf{a}_i^q - \mathbf{d}_i^q) f'(n_i^q)$ e $n_i^q = \sum_{k=0}^R w_{ik}^q(u) \mathbf{p}_k^q$ avremo

$$\Delta w_{ik}^q = -\eta \delta_i^q p_k^q \text{ nota come } \delta rule.$$

L'innovazione principale di tale nuovo modello di ANN fu l'introduzione di pesi differenti per differenti connessioni, il cui aggiornamento permette alle reti di apprendere da nuovi elementi forniti.

L'incapacità di tali rete nel discriminare insiemi di dati i cui elementi non sono linearmente separabili fu messa in evidenza nel 1969 dai matematici Minsky e Papert [23], nel loro libro Perceptrons, e portò all'ideazione di una nuova tipologia di reti, basata sui concetti delle reti perceptrone, organizzate tra loro in strati connessi, ovvero le reti multistrato [22, 24].

2.2.1 Reti multistrato

Le reti multistrato rappresentano una generalizzazione delle reti perceptrone. Esse presentano più strati di neuroni, solitamente denominati strato di input, strato nascosto e strato di output, in cui esiste un verso di connessione tra i neuroni di uno strato all'altro (Figura 2.21) e spesso ci si riferisce ad uno strato nascosto anche se esso può identificare più strati nascosti. I valori di output di ogni strato, escluso l'ultimo, rappresentano i valori di ingresso dello strato successivo, il propagarsi dei valori da uno strato al successivo è il motivo per cui tali reti vengono denominate reti neurali feed forward (Multi Layer Feed Forward Neural Network, MLFFNN), in cui il segnale viene propagato dallo strato di input a quello di output. I neuroni appartenenti al medesimo strato non sono comunicanti tra loro, bensì saranno collegati ai neuroni dello strato successivo. In genere gli strati interni vengono denominati nascosti poiché non hanno nessuna interazione con l'esterno, diversamente dallo strato di input e di output; sono proprio gli strati nascosti che permettono di apprendere le caratteristiche

di insiemi di dati non linearmente separabili, incrementando le capacità di computazionali delle reti neurali. L'introduzione di un livello (layer) incrementa notevolmente le capacità di computazione delle reti, infatti con uno strato è possibile individuare un semipiano, con due strati è possibile individuare l'intersezione di semipiani, quindi regioni convesse, con più strati è possibile individuare intersezioni di regioni convesse e quindi figure arbitrarie, con le quali è possibile delimitare classi di elementi complesse e connesse (Figura 2.22).

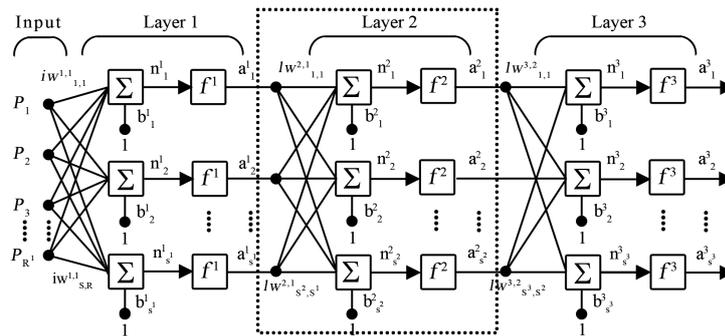


Figura 2.21: Rappresentazione di una rete MLFFNN.

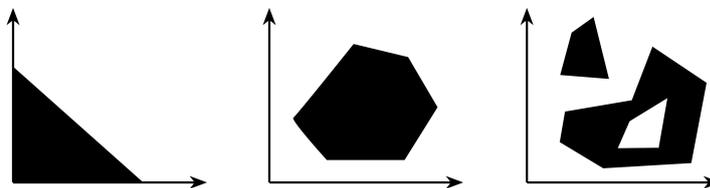


Figura 2.22: Usando la funzione di attivazione a gradino, a sinistra la regione delimitata da un perceptrone, al centro la regione delimitata da rete neurale multistrato a 1 livello hidden, a destra la regione delimitata da rete neurale multistrato con 2 livelli hidden che può delimitare aree arbitrarie.

In riferimento alla Figura 2.21, una generica rete MLFFNN sarà caratterizzata da un numero L di strati, ed ciascuno di essi avrà un numero $S^{(\ell)}$ di neuroni; con a_j^ℓ indichiamo l'output del j -esimo neurone del layer ℓ , con $\ell \in \{1 \dots L\}$.

Consideriamo una rete feed forward avente un numero arbitrario di strati nascosti, con funzioni di attivazione non lineari; il generico output dell'unità j -esima del layer $\ell + 1$ sarà

ottenuto dal valore di f di $n_j^{\ell+1}$, esprimibile come:

$$\mathbf{a}_j^{\ell+1} = \mathbf{f}_j^{\ell+1}(\mathbf{n}_j^{\ell+1}) = \mathbf{f}_j^{\ell+1} \left(\sum_{k=1}^{S^\ell} LW_{jk}^{\ell+1,\ell} + \mathbf{a}_k^\ell + \mathbf{b}_k^{\ell+1} \right)$$

con $j = 1, \dots, S^{\ell+1}$ e dove $LW_{jk}^{\ell+1,\ell}$ indica il peso di input per il j -esimo elemento (unità) del layer $\ell + 1$ dalla k -esima unità del precedente layer ℓ . Nel caso specifico in cui ℓ rappresenta l'ultimo layer, strato di output, quindi per $\ell = L$ avremo, in notazione matriciale:

$$\mathbf{a}^L = \mathbf{f}^L(LW^{L,L-1} \mathbf{f}^{L-1}(\dots \mathbf{f}^i(\dots \mathbf{f}^1(IW^{1,1}p + b_1) + \dots + b_i) + b_{L-2}) + b_{L-1}) + b_L) \quad (2.2.8)$$

Considerando un tipo di apprendimento supervisionato, poiché conosciamo soltanto l'output atteso dell'ultimo strato, non possiamo calcolare direttamente l'errore degli strati intermedi, quindi non è permesso in alcun modo identificare, in caso di errore, le unità interne responsabili, e di conseguenza, quale peso modificare. Uno dei primi algoritmi per l'aggiornamento dei pesi delle reti feedforward multilayer è l'algoritmo *back propagation*, il quale utilizza il metodo di retropropagazione dell'errore, dagli strati di output a quelli intermedi, i quali a loro volta aggiorneranno i propri pesi w ; la strategia di base è simile a quella del metodo del gradiente discendente.

Consideriamo l'errore \mathbf{E} per il generico input q , $q = 1, \dots, M$; E^q , similmente alla equazione (2.2.6) dipenderà dai pesi w_{ik} , l'aggiornamento dei pesi nel caso dell'apprendimento incrementale sono espressi dalle seguenti relazioni:

$$\mathbf{E}^q(u) = \frac{1}{2} \sum_{i=1}^{S^\ell} (\mathbf{a}_{\ell i}^q - \mathbf{d}_i^q)^2$$

$$\Delta w_{i,k}^q = -\eta \frac{\partial \mathbf{E}^q}{\partial w_{ik}}$$

considerando che è possibile applicare *la regola della catena* per le derivate parziali e ottenere:

$$\frac{\partial \mathbf{E}^q}{\partial w_{ik}} = \frac{\partial \mathbf{E}^q}{\partial n_{\ell i}^q} \frac{\partial n_{\ell i}^q}{\partial w_{ik}} = \delta_i^q \frac{\partial n_{\ell i}^q}{\partial w_{ik}} \quad (2.2.9)$$

Ponendo

$$\delta_i^q = \frac{\partial \mathbf{E}^q}{\partial n_{\ell i}^q} \quad (2.2.10)$$

e applicando le seguenti sostituzioni

$$n_{\ell j}^q = \sum_{k=0}^{S^\ell} w_{jk} a_{(\ell-1)k}^q \quad \text{con} \quad w_{j0} = b_j, \quad a_{\ell 0}^q = 1 \quad \forall \ell, q$$

in analogia alla δ rule,

$$\Delta w_{ik}^q(u) = -\eta \delta_i^q a_{(\ell-1)k}^q$$

dove $i = 1, \dots, S^\ell$, $k = 0, \dots, S^{\ell-1}$ e $\ell = 1, \dots, L$ permettendoci di calcolare la δ rule per i neuroni a livello ℓ a partire dal livello $\ell - 1$. Considerando l'ultimo layer, quindi per $\ell = L$ avremo:

$$\begin{aligned} \frac{\partial \mathbf{E}^q}{\partial w_{i,k}} &= \frac{\partial \frac{1}{2} \sum_{j=1}^{S^L} (\mathbf{a}_{Lj}^q - \mathbf{d}_j^q)^2}{\partial w_{ir}} = \frac{\partial \frac{1}{2} \sum_{j=1}^{S^{L-1}} \left(f \left(\sum_{r=0}^{S^{L-1}} w_{jr}^q \mathbf{a}_{(L-1)r}^q \right) - \mathbf{d}_i^q \right)^2}{\partial w_{ik}} = \\ &= (\mathbf{a}_{Li}^q - \mathbf{d}_i^q) \frac{\partial f \left(\sum_{r=0}^{S^{L-1}} w_{ir}^q \mathbf{a}_{(L-1)r}^q \right)}{\partial w_{ik}} = (\mathbf{a}_{Li}^q - \mathbf{d}_i^q) \frac{f'(n_{Li}^q)}{\partial w_{ik}} = \\ &= (\mathbf{a}_{Li}^q - \mathbf{d}_i^q) \frac{f'(n_{Li}^q)}{\partial n_{Li}^q} \frac{\partial n_{Li}^q}{\partial w_{ik}} = (\mathbf{a}_{Li}^q - \mathbf{d}_i^q) f'(n_{Li}^q) \mathbf{a}_{(L-1)k}^q \end{aligned} \quad (2.2.11)$$

Utilizzando la *regola della catena* e sostituendo con (2.2.10):

$$\frac{\partial \mathbf{E}^q}{\partial w_{i,k}} = \frac{\partial \mathbf{E}^q}{\partial n_{\ell i}^q} \frac{\partial n_{\ell i}^q}{\partial w_{i,k}} = \delta_i^q \frac{\partial n_{\ell i}^q}{\partial w_{i,k}} = \delta_i^q \mathbf{a}_{(L-1)k}^q \quad (2.2.12)$$

Uguagliando (2.2.11) con (2.2.12) si ottiene la δ rule per i neuroni del livello L

$$\delta_i^q = (\mathbf{a}_{Li}^q - \mathbf{d}_i^q)^2 f'(n_{Li}^q)$$

Per il generico livello ℓ si avrà:

$$\delta_i^q = \frac{\partial \mathbf{E}^q}{\partial n_{\ell i}^q} = \sum_{j=1}^{S^{\ell+1}} \frac{\partial \mathbf{E}^q}{\partial n_{(\ell+1)j}} \frac{\partial n_{(\ell+1)j}}{\partial n_{\ell i}^q} \quad (2.2.13)$$

Poichè

$$\frac{\partial n_{(\ell+1)j}}{\partial n_{\ell i}^q} = \frac{\partial \sum_{k=0}^{S^\ell} w_{jk}^q \mathbf{a}_{\ell k}^q}{\partial n_{\ell i}^q} = \frac{\sum_{k=0}^{S^\ell} \partial \left(w_{jk}^q \mathbf{a}_{\ell k}^q \right)}{\partial n_{\ell i}^q} = \frac{\sum_{k=0}^{S^\ell} w_{jk}^q \partial \mathbf{a}_{\ell k}^q}{\partial n_{\ell i}^q} = w_{ji} f'(n_{\ell i}^q) \quad (2.2.14)$$

sostituendo in (2.2.13) il valore $\frac{\partial n_{(\ell+1)j}}{\partial n_{\ell i}}$ si ottiene:

$$\delta_i^q = \sum_{j=1}^{S^{\ell+1}} \frac{\partial \mathbf{E}^q}{\partial n_{(\ell+1)j}} f'(n_{\ell i}^q) w_{ji} \quad (2.2.15)$$

Infine eguagliando (2.2.10) con (2.2.15) si ottiene la formulazione della *δrule* per il livello ℓ a partire dal livello $\ell + 1$

$$\delta_i^q = \sum_{j=1}^{S^{\ell+1}} \delta_j^q f'(n_{\ell i}^q) w_{ji} = f'(n_{\ell i}^q) \sum_{j=1}^{S^{\ell+1}} w_{ji} \delta_j^q$$

con $i = 1, \dots, S^\ell$

Possiamo affermare che applicare l'algoritmo della back-propagation equivale a trovare la matrice ottimale dei pesi che minimizza la funzione di errore:

$$\mathbf{W}^* = \arg \min_{\mathbf{w}} \mathbf{E}(\mathbf{W})$$

Un punto debole dell'apprendimento basato sull'algoritmo backpropagation sta nella possibilità di raggiungere un punto di minimo locale della funzione \mathbf{E} ; questo potrebbe accadere perché, mentre per una rete con un unico layer, la funzione di errore avrà una forma convessa e quindi con un solo punto di minimo(assoluto), per una rete con più layers la funzione di errore è una funzione generica non lineare dei pesi, la quale potrebbe presentare punti di minimo locale o di flesso (Figura 2.23).

L'apprendimento di una MLFFNN basato sul principio di *backpropagation* può essere rappresentato dal seguente algoritmo:

Sia \mathbf{M} il numero di pattern (elementi) del training set

- Inizializzare la matrice dei pesi W_{ik} con valori randomici, per ogni livello

FOR q=1:M

- Propaga in avanti il pattern p^q nei livelli della rete,(ogni livello propaga al successivo il proprio output) $\mathbf{a}^l = \mathbf{f}'(\mathbf{n}^\ell) = \sum_{k=0}^{S^{\ell-1}} \mathbf{W}_{\ell k} \mathbf{a}_{(\ell-1)k}$
- Calcola l'output della rete a^L e l'errore $\mathbf{E}^q(W)$
- Calcola $\delta_i^q = (\mathbf{a}_{Li}^q - \mathbf{d}_{Li}^q)^2 f'(n_{Li}^q)$ con $i \in S^L$

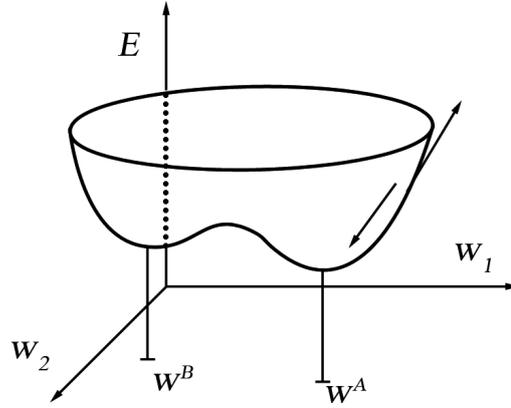


Figura 2.23: Rappresentazione geometrica della funzione di Errore $E(w)$ come una superficie nello spazio dei pesi. Il punto W^A indica il minimo globale, mentre W^B è punto di minimo locale ∇E indica il vettore gradiente della funzione di errore.

- Per ciascun livello $\ell = L \dots 1$ calcola $\delta_i^q = f'(n_{Li}^q) \sum_{k=1}^{S^\ell} w_{kj} \delta_k^q \forall j = 0 \dots S^{\ell-1}, \forall k \in S^\ell$
- Per ciascun livello $\ell = L \dots 1$ aggiorna i pesi:

$$w_{kj}^{q+1} = w_{kj}^q + \Delta w_{kj}^q = w_{kj}^q - \eta \delta_i^q a_{(\ell-1)j}^q \forall j = 0 \dots S^{\ell-1}, \forall k \in S^\ell$$

END

Lo scopo dell'addestramento quindi non è quello di semplice interpolazione dei dati forniti in input, piuttosto quello di modellare il processo che ha generato i dati, motivo per cui la selezione dei dati da utilizzare nel processo di apprendimento, la scelta della funzione di errore e la struttura della rete giocano un ruolo importante. Nonostante esistano in letteratura lavori [25, 26] che fornisco delle stime per la scelta della cardinalità dell'insieme di apprendimento (training set), spesso si seguono approcci di tipo sperimentale, ovvero addestrando diverse reti neurali con configurazioni diverse, con differenti inizializzazioni dei pesi, e confrontando le prestazioni delle reti utilizzando tecniche di cross-validation.

Tali tecniche prevedono un partizionamento dei dati di input, una parte di essi viene utilizzata per addestrare la rete, l'altra per testarle. L'inizializzazione dei pesi, gioca un ruolo fondamentale anche nel caso delle reti feedforward multistrato. In genere, si predilige un'inizializzazione random con piccoli valori dei pesi per le diverse reti addestrate, talvolta i pesi vengono perturbati per evitare minimi locali. Il parametro di learning η è un valore

che è possibile gestire al fine di evitare l'oscillazione attorno ad un punto di minimo locale o per definire un apprendimento proporzionale al numero di epoche; inoltre viene introdotto anche un termine detto momentum e identificato con α , con valore compreso tra 0 e 1, che mantiene memoria dell'ultimo aggiornamento. Quando il gradiente mantiene la stessa direzione, aumenterà le dimensioni degli step compiuti in direzione del minimo. Qualora si utilizza un momentum prossimo a 1, sarà necessario ridurre il tasso globale di apprendimento. Combinando un alto tasso di apprendimento con un alto valore del momentum, si tenderà al minimo.

$$\Delta w_{ik}(u) = -\eta \frac{\partial E}{\partial w_{ik}} + \alpha \Delta w_{ik}(u-1)$$

Per rendere più robusto l'apprendimento, prima di aggiornare i pesi si valuta se con l'aggiornamento dei pesi l'errore cresce o diminuisce:

$$\Delta\eta = \begin{cases} +\rho\eta(u-1) & \text{se } \Delta E(W, z) < 0 \text{ per } z = u-1, \dots, u-U \\ -\sigma\eta(u-1) & \text{se } \Delta E(W, z) > 0 \text{ per } z = u-1, \dots, u-U \\ 0 & \text{altrimenti} \end{cases} \quad (2.2.16)$$

dove

$$\Delta E(W, u) = E(W, u) - E(W, u-1)$$

e

$$\eta(u) = \eta(u-1) + \Delta\eta(u)$$

Durante la fase di apprendimento di una rete neurale potrebbe verificarsi il fenomeno dell'overfitting, ovvero quando il modello appreso dalla rete tende a interpolare i dati di training, a svantaggio della generalizzazione, deducibile dal valore assunto dall'errore che tenderà a zero. Una tecnica per gestire tale fenomeno, denominata early stopping, consiste nel considerare un insieme di validazione ausiliario e monitorare il valore dell'errore su tale insieme; indicato con $E_T(t)$ l'errore sull training set e $E_V(t)$ l'errore sull'insieme di validazione, $E_V(t)$

inizialmente diminuirà in accordo a $E_T(t)$, ma con il procedere delle interazioni $E_V(t)$ potrebbe cominciare ad aumentare, in questo caso l'apprendimento è interrotto prematuramente e come insieme dei pesi sarà scelto quello per cui si ha il minimo di $E_v(t)$ (Figura 2.24).

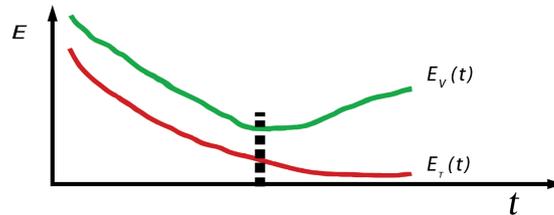


Figura 2.24: Grafico dell'errore sul training set e validation set per determinare l'interruzione dell'apprendimento per evitare l'overfitting.

Come discusso precedentemente, anche il numero di layer nascosti e il numero di neuroni con cui la rete neurale è costruita determina la capacità di apprendimento di una rete. Non esiste una regola generale che permetta di determinare in modo preciso il numero di layer e neuroni che necessita una rete per la modellazione del processo che ha generato i dati, esistono però delle strategie che ne permettono una valutazione: algoritmi di growing e pruning. Nell'algoritmo growing inizialmente si addestra una piccola rete e si aggiungono nuovi neuroni durante il processo di training valutando di volta in volta se le performance migliorano. Il procedimento inverso è la strategia di base dell'algoritmo pruning, in cui la rete inizialmente ha un numero sufficientemente grande di neuroni, molti dei quali saranno eliminati durante la fase di training, eliminando le unità i cui pesi sono prossimi a zero. Le unità possono essere rimosse secondo due modalità: sensitivity e penalty-term. Con il primo metodo, ogni neurone, dopo la fase di apprendimento, viene temporaneamente rimosso e si valuta la variazione della funzione di errore. Nel secondo metodo, la funzione errore, cfr. (2.2.17), viene modificata opportunamente, attraverso il parametro β , in modo che i pesi delle unità non necessarie, dopo la fase di apprendimento, tenderanno a zero.

$$E'(\mathbf{W}) = E(\mathbf{W}) + \beta \|\mathbf{W}\| \quad (2.2.17)$$

2.2.2 Reti neurali e problemi di classificazione

Le reti neurali possono essere utilizzate anche per risolvere problemi di classificazione, tramite un modello discriminante non lineare o addirittura modellare le probabilità a posteriori di appartenenza alle classi considerate. Quando una rete neurale modella le probabilità a posteriori, il numero di neuroni nello stato di output è pari al numero di classi rappresentate e l'attivazione dell'unità k rappresenta la probabilità a posteriori per la classe k . La fase di addestramento di una rete neurale probabilistica, assumendo che abbia L strati, corrisponde all'apprendimento, a partire da un training set $X = \{(x_1, t_1), (x_2, t_2), \dots, (x_M, t_M)\}$ dei valori dei coefficienti \mathbf{W} . L'apprendimento viene effettuato minimizzando una funzione di errore dipendente dalle ipotesi probabilistiche poste, in funzione del problema considerato: tale funzione di errore, nel criterio di maximum likelihood, è tale che la sua minimizzazione equivale alla massimizzazione della verosimiglianza. Nel caso di classificazione binaria, l'unità di output ha funzione di attivazione sigmoide, $f(n) = \frac{1}{1+e^{-n}}$. Posto $y(x, \mathbf{W}) = p(C1|x)$, la probabilità condizionata del target può essere espressa dalla distribuzione di Bernoulli:

$$p(t|\mathbf{x}, \mathbf{W}) = y(\mathbf{x}, \mathbf{W})^t (1 - y(\mathbf{x}, \mathbf{W}))^{1-t}$$

e, sull'intero training set X , la verosimiglianza sarà

$$L(t|\mathbf{X}, \mathbf{W}) = \prod_{i=1}^M y(\mathbf{x}_i, \mathbf{W})^{t_i} (1 - y(\mathbf{x}_i, \mathbf{W}))^{1-t_i}$$

e la log-verosimiglianza, denominata cross-entropy è

$$l(t|\mathbf{X}, \mathbf{W}) = \sum_{i=1}^M (t_i \ln y_i + (1 - t_i) \ln(1 - y_i))$$

La funzione di errore sarà $E(\mathbf{W}) = -l(t|\mathbf{X}, \mathbf{W})$ e, per quanto riguarda la derivata rispetto a $a_{L,k}$ abbiamo che

$$\frac{\partial E(\mathbf{W})}{\partial a_{L,k}} = -t_k \frac{1}{y_k} \frac{\partial y_k}{\partial a_{Lk}} + (1 - t_k) \frac{1}{1 - y_k} \frac{\partial y_k}{\partial a_{Lk}}$$

e, dato che $y_k = f(a_{Lk})$, allora

$$\frac{\partial y_k}{\partial a_{Lk}} = \frac{f'(a_{Lk})}{1 - f(a_{Lk})} = y_k(1 - y_k)$$

Pertanto concludiamo

$$\frac{\partial E(\mathbf{W})}{a_{Lk}} = -t_k \frac{1}{y_k} (1 - y_k) + (1 - t_k) \frac{1}{1 - y_k} y_k (1 - y_k) = -t_k (1 - y_k) + (1 - t_k) y_k = y_k - t_k$$

Nel caso in cui $K > 2$, la funzione di attivazione sigmoide è sostituita da una sua generalizzazione nota come softmax, o esponenziale normalizzato:

$$y_k = \frac{\exp(a_{Lk})}{\sum_{j=1}^K \exp(a_{Lj})}$$

per cui possiamo esprimere le probabilità a posteriori delle K classi per l'input x come:

$$p(C_k|x) = \frac{p(x|C_k)p(C_k)}{\sum_{j=1}^K p(x|C_j)p(C_j)} = \frac{\exp(a_{Lk})}{\sum_{j=1}^K \exp(a_{Lj})}$$

con $a_k(x) = \ln(p(x|C_j)p(C_j))$.

La funzione di errore in questo caso, come visto in precedenza, è $E(W) = -\sum_{j=1}^K t_j \log y_j$ per cui:

$$\frac{\partial E(\mathbf{W})}{a_{Lk}} = -\frac{\partial}{\partial a_{Lk}} \sum_{j=1}^K t_j \log y_j = -\sum_{j=1}^K \frac{t_j}{y_j} \frac{\partial y_j}{\partial a_{Lk}} = \sum_{j \neq k} t_j y_k - t_k (1 - y_k) = y_k \sum_{j=1}^K t_j - t_k = y_k - t_k$$

in quanto $\sum_{j=1}^K t_j = 1$.

Le reti neurali sono dunque uno strumento versatile, adattabile a diversi contesti e scopi computazionali. In letteratura esistono diverse tipologie di reti neurali oltre a quelle feed forward, per una trattazione approfondita si rimanda a [22, 24].

2.3 Trasformata simmetria

Giustificata dalla morfologia delle componenti da noi qui analizzate, abbiamo concentrato il nostro interesse sulle metodiche atte all'individuazione dei punti di interesse che identificano o delimitano un oggetto, come nella ricerca dei corner, o punti all'interno degli oggetti, come i centri di simmetria. L'obiettivo consiste quindi nel riconoscere la presenza e il numero delle singole componenti in un'immagine, le cui caratteristiche sono in accordo con le proprietà espresse dai punti di interesse. Nello specifico siamo interessati all'individuazione dei punti che evidenziano una forte simmetria radiale.

Diversi lavori sono stati presentati in letteratura sull'individuazione dei centri di simmetria radiale: alcuni riguardano il tracking di oggetti, il riconoscimento dei volti e analisi di immagini astronomiche. La nostra scelta è ricaduta sul lavoro di Gareth Loy e Alexander Zelinsky [27] di cui tracciamo adesso le principali proprietà.

Trasformata veloce per la simmetria radiale

In [27] è proposto un nuovo operatore basato sul gradiente che individua efficientemente in immagini digitali i punti ad alta simmetria radiale. L'approccio si ispira a quelli sulla trasformata simmetria generalizzata [28, 29, 30] a quelli sulla similarità [31] e sulla trasformata di Hough generalizzata [32, 33]. La trasformata determina il contributo che ogni pixel fornisce ai pixel simmetrici attorno ad esso, diversamente dalle precedenti tecniche che consideravano il contributo che i pixel di un intorno fornivano al pixel centrale. Un parametro che la trasformata richiede in input è l'insieme N dei raggi entro i quali sono ricercate le caratteristiche di simmetria radiale. Il valore della trasformata in un dato punto per $n \in N$ indica il contributo del gradiente alla simmetria radiale a distanza n da quel punto.

L'algoritmo è qui di seguito sintetizzato considerando in input una immagine I a scala di grigio:

- Determinare il gradiente g dell'immagine in input.
- Per ogni raggio considerato:

- Per ciascun elemento gradiente:
 Determinare i pixel interessati.
 Calcolare l'immagine proiezione dell'orientamento O_n e l'immagine proiezione magnitudo M_n .
 Combinare O_n e M_n per creare i contributi simmetria non filtrati F_n .
 Calcolare l'immagine simmetria per il raggio n , S_n dall'immagine F_n convoluta con A_n .
- Sommare S_n per tutti i raggi $n \in N$ per determinare l'immagine simmetria finale S .

L'immagine gradiente g di I è ottenuta applicando il classico operatore di Sobel [34]:

$$K = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} \quad \text{e} \quad K^T = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}.$$

Per ogni raggio considerato si determinano i pixel interessati esaminando il gradiente. Ciascun elemento gradiente diverso da zero genera un *pixel interessato positivamente* e un *pixel interessato negativamente*. Un pixel interessato positivamente è definito come quel pixel a cui il vettore gradiente punta con una distanza n dal pixel considerato; analogamente un pixel interessato negativamente è definito come quel pixel a distanza n dal pixel considerato, ma in direzione opposta al vettore gradiente (Figura 2.25).

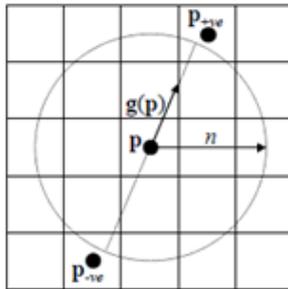


Figura 2.25: Individuazione dei punti interessati positivamente e negativamente.

Le coordinate del pixel interessato positivamente e del pixel interessato negativamente sono date quindi:

$$\mathbf{p}_{+ve}(\mathbf{p}) = \mathbf{p} + \left(\frac{n \mathbf{g}(\mathbf{p})}{\|\mathbf{g}(\mathbf{p})\|} \right) \quad \text{e} \quad \mathbf{p}_{-ve}(\mathbf{p}) = \mathbf{p} - \left(\frac{n \mathbf{g}(\mathbf{p})}{\|\mathbf{g}(\mathbf{p})\|} \right)$$

Le immagini O_n e M_n sono accumulatori inizialmente nulli che vengono incrementati rispettivamente di 1 e $\|\mathbf{g}(\mathbf{p})\|$ nei punti $\mathbf{p}_{+ve}(\mathbf{p})$. Analogamente nei punti $\mathbf{p}_{-ve}(\mathbf{p})$ gli accumulatori O_n e M_n sono decrementati di 1 e $\|\mathbf{g}(\mathbf{p})\|$.

$$O_n(\mathbf{p}_{+ve}(\mathbf{p})) = O_n(\mathbf{p}_{+ve}(\mathbf{p})) + 1 \quad \text{e} \quad O_n(\mathbf{p}_{-ve}(\mathbf{p})) = O_n(\mathbf{p}_{-ve}(\mathbf{p})) - 1$$

$$M_n(\mathbf{p}_{+ve}(\mathbf{p})) = M_n(\mathbf{p}_{+ve}(\mathbf{p})) + \|\mathbf{g}(\mathbf{p})\| \quad \text{e} \quad M_n(\mathbf{p}_{-ve}(\mathbf{p})) = M_n(\mathbf{p}_{-ve}(\mathbf{p})) - \|\mathbf{g}(\mathbf{p})\|$$

Il contributo della simmetria radiale per un raggio n è definita come la convoluzione :

$$S_n = F_n * A_n$$

dove:

$$F_n = \|\tilde{O}_n(\mathbf{p})\|^\alpha \tilde{M}_n(\mathbf{p})$$

$$\tilde{O}_n(\mathbf{p}) = \frac{O_n}{\max_p \{\|O_n(\mathbf{p})\|\}}$$

$$\tilde{M}_n(\mathbf{p}) = \frac{M_n}{\max_p \{\|M_n(\mathbf{p})\|\}}$$

A_n è un kernel Gaussiano, per diffondere l'influenza dei pixel $\mathbf{p}_{+ve}(\mathbf{p})$ o $\mathbf{p}_{-ve}(\mathbf{p})$ in funzione del raggio n . È stato scelto questo kernel poichè è radialmente simmetrico e separabile, rendendo la convoluzione più efficiente (Figura 2.26).

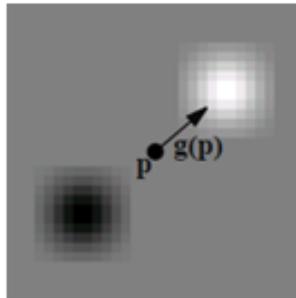


Figura 2.26: Schematizzazione dei contributi dei pixel considerati positivamente e negativamente.

La trasformata si completa sommando tutti i contributi simmetria S_n considerati:

$$S = \sum_{n \in N} S_n$$

Se il gradiente è calcolato per i punti di transizione da scuro a chiaro, allora nell'immagine della trasformata simmetria avremo valori positivi per le regioni chiare ad alta simmetria radiale e valori negativi per regioni scure. Ovviamente considerando i soli pixel $\mathbf{p}_{+ve}(\mathbf{p})$ o $\mathbf{p}_{-ve}(\mathbf{p})$ possiamo applicare la trasformata per esaltare selettivamente le regioni ad alta simmetria radiale chiare o scure.

Capitolo 3

Metodologia

3.1 Stato dell'arte

Nel caso specifico delle metodologie orientate all'individuazione di patologie, Cicero et al. [35] hanno mostrato che la misurazione della traslucenza nucale richiede elevate competenze la cui acquisizione è tutt'altro che banale. Tra i più rilevanti lavori descritti in letteratura su questo argomento, il primo è cronologicamente dovuto a Bernardino et al. [36] in cui è descritto il miglioramento del contrasto attraverso l'applicazione dell'equalizzazione dell'istogramma della luminosità; l'utente identificherà manualmente un punto per ciascun bordo che delimita la traslucenza e da questi i bordi saranno tracciati automaticamente tramite i filtri di Sobel e Canny i quali risultano non particolarmente efficaci su immagini ecografiche, poichè la presenza di rumore e regioni poco omogenee sono tali da ottenere bordi irregolari, frastagliati e sensibili al rumore.

In Lee et al. [37] è proposta una metodologia semiautomatica basata sulla programmazione dinamica, su immagini in cui è stato applicato un filtro anisotropo non lineare per ridurre lo speckle noise [38]. La metodologia viene applicata su trenta immagini delle quali è noto il valore della misura effettuata manualmente da un medico; i risultati delle misure manuali e/o semiautomatiche riportano i valori di media e deviazione standard espressi in millimetri e pari rispettivamente a 2.79 ± 0.77 e 2.81 ± 0.79 . Il metodo proposto tende a sottostimare lo spessore della traslucenza nucale. Inoltre la posizione fetale limita ulteriormente l'automatismo della metodologia che può essere applicata in immagini in cui il feto assume una

posizione orizzontale.

In Catanzariti et al. [39] è stata proposta una metodologia basata sulla programmazione dinamica, in cui la funzione costo per la segmentazione dei bordi che delimitano la translucenza è stata notevolmente migliorata, non essendo vincolata ad alcun parametro, e il processo è applicabile senza l'intervento dell'utente che non dovrà più selezionare i punti iniziali per l'individuazione dei bordi. Una analisi qualitativa mostra l'efficienza del metodo.

In Nirmala et al. [40] le immagini sono pre-elaborate applicando un filtro mediano, per rimuovere il rumore speckle. L'individuazione della regione in cui risiede la NT è affidata all'utente, su tale regione viene applicato l'algoritmo *mean shift* per la segmentazione. All'immagine segmentata, successivamente, è applicato l'operatore Canny per l'estrazione di quei bordi che delimitano la translucenza nucale. La *blob analysis* viene proposta per la misurazione dello spessore della translucenza. I valori medi degli spessori delle translucenze nicali sono suddivisi tra l'undicesima e la quattordicesima settimana di gestazione e calcolati su circa dieci soggetti per settimana. Gli autori riportano un confronto quantitativo tra i valori medi degli spessori della translucenza considerati nella norma e non.

In Deng et al. [41] viene proposto uno schema semi-automatico: le immagini vengono pre-processate con un filtro morfologico per ridurre il rumore presente e successivamente sogliate con un valore calcolato empiricamente. L'utente seleziona due punti iniziali e su di essi, mediante *gradient vector flow snake*, è individuato il bordo; successivamente il bordo ottenuto è migliorato tramite un algoritmo di programmazione dinamica. Infine lo spessore e l'area della translucenza nucale sono calcolati. È riportato un confronto qualitativo dei risultati ottenuti su immagini sintetiche e reali.

Un modello gerarchico per l'identificazione automatica della regione nucale è proposto in [42]: tre classificatori support vector machine (SVM) sono addestrati per rappresentare la regione nucale, la testa e il corpo del feto. Gli esperimenti sono stati condotti su 345 immagini e sono state considerate come corrette quelle regioni la cui area copre almeno il 50% dell'area selezionata dall'esperto; i risultati sono mostrati per diverse cardinalità del training set e si osserva che, nel caso in cui la dimensione del training set è pari a 170 (valore massimo riportato), si ottengono delle percentuali di riconoscimento pari a circa il 100%, 97% e 55% rispettivamente per le regioni del corpo, della testa e della translucenza nucale.

In Deng et al. [43] la medesima metodologia è stata rivisitata aggiungendo un ulteriore livello

rappresentativo del profilo fetale (sinistro o destro), grazie al quale le prestazioni sono migliorate: gli esperimenti sono stati condotti su 690 immagini e, utilizzando il 50% del training set per l'addestramento dei classificatori, si ottiene una percentuale di riconoscimento del 93% per la regione nucale.

Sebbene i dettagli tecnici non siano stati divulgati, in Moratalla et al. [44] viene presentato l'ambiente SonoNT, già integrato e commercializzato in alcuni ecografi, che permette una misurazione semi-automatica della translucenza nucale; l'utente racchiude la translucenza nucale in un box e il posizionamento deve essere tale che il box contenga lo spessore massimo della NT. Il tool traccia i bordi inferiore e superiore utilizzando le informazioni del gradiente e della luminosità all'interno del riquadro; infine identifica la massima distanza verticale tra i due contorni. Il data set su cui sono stati condotti gli esperimenti è composto da 48 immagini, ottenuto da una disposizione casuale di 12 immagini iniziali; 20 sono gli esperti che hanno effettuato le misurazioni; le valutazioni quantitative sono state condotte sulla variabilità delle misurazioni inter-operator mostrando che nel caso delle misure manuali si ha una variabilità di circa 0.20mm, che si riduce a 0.01mm nel caso di misure eseguite tramite SonoNT. Va sottolineato che in questo caso la misura è sovrastimata di circa 0.40mm.

Un metodo automatico per l'individuazione e la misurazione della translucenza nucale è presentato in Supriyanto et al. [45]: la regione contenente la translucenza nucale è identificata con l'uso di una rete neurale multistrato che riceve in input sottocampioni di immagini restituendo in output la percentuale di appartenenza alla classe rappresentativa per la translucenza. Una volta identificati i punti con un'alta probabilità di appartenenza alla regione nucale, la translucenza è identificata e su di essa saranno estratti i bordi con un algoritmo automatico basato su misure di intensità, conosciuto come "bidirectional iterations forward propagations method". I risultati sono stati ottenuti su 30 immagini per le quali è stato riportato un indice di correlazione cumulativo sull'intero data set pari a 0.98, tra le misurazioni automatiche e manuali.

Nel 2013 Park et al. [46] hanno proposto una metodologia automatica per la misurazione della translucenza, in cui dopo aver individuato la posizione della testa, viene individuata la regione nucale, basandosi su relazioni statistiche che intercorrono tra esse. L'algoritmo di Dijkstra per il calcolo del cammino più corto è applicato per individuare i bordi interno ed esterno delle due componenti che delimitano la translucenza; all'interno di queste regioni sono

identificati i semi che verranno utilizzati dall'algoritmo di segmentazione graph cut, ed infine sarà misurato il diametro della translucenza nucale. I risultati per l'individuazione delle due regioni sono stati condotti su un dataset di 196 immagini e sono riportate le differenze tra gli angoli di inclinazione, i centri e le dimensioni delle regioni individuate automaticamente e manualmente, relativamente alla testa e alla translucenza nucale. Valutazioni qualitative per l'estrazioni dei bordi e della misura della translucenza sono riportati graficamente, mostrando i cinque migliori e peggiori casi analizzati.

La metodologia semiautomatica proposta da Sonia et al. [47] ha l'obiettivo di classificare tramite SVM le immagini che presentano uno spessore della translucenza nucale nella norma e quelle il cui spessore è superiore alla norma. L'estrazione delle caratteristiche è eseguita applicando una trasformata wavelet discreta. Nella metodologia lo spessore non è quantitativamente misurato, ovvero non vi è né l'estrazione dei bordi né la misura del diametro massimo. Gli esperimenti sono stati condotti su 71 immagini e sono stati ottenuti il 93.8% e 74.4% di sensibilità e specificità, rispettivamente.

I lavori riportati precedentemente focalizzano l'attenzione sulla individuazione e la misurazione della translucenza nucale, lasciando all'esperto il compito di selezionare la corretta sezione sagittale mediana, condizione necessaria affinché possa essere eseguita una corretta misurazione, e richiedendo talvolta il successivo posizionamento manuale del box contenente la regione di interesse. Al meglio delle nostre conoscenze sono pochi i lavori presenti in letteratura che propongono metodologie per l'individuazione di una sezione sagittale mediana acquisita con modalità bidimensionale classica.

Uno studio orientato all'individuazione automatica della sezione sagittale mediana e alla misurazione della translucenza è riportato in Anzalone et al. [48]. La metodologia è composta da due fasi: inizialmente l'attenzione è rivolta al plesso corioideo e all'osso mascellare per stabilire se la sezione sagittale è mediana; l'algoritmo di Perona e Malik [49] riduce la presenza di rumore speckle e successivamente la trasformata di Hough è applicata per analizzare il plesso; infine un algoritmo di region growing individua l'osso mascellare. Nella seconda fase, con l'identificazione del mento attraverso template matching, viene determinato il profilo fetale utile per l'individuazione della regione nucale, la quale avviene nuovamente tramite template matching; infine la procedura proposta in [39] è applicata per l'estrazione dei bordi. L'efficienza del metodo per il riconoscimento della sezione sagittale mediana viene valutata su

171 immagini con un corretto riconoscimento del 92%. Sono state condotte sperimentazioni per determinare la migliore metrica per il template matching e per stabilire il numero di template necessari per l'identificazione del mento e della translucenza. Dai risultati ottenuti si evince che la migliore metrica è la *mutual information*, confrontata con la *normalised cross correlation* e con la *sum of spread difference*, mentre il numero ottimale di template da utilizzare è 3 in entrambi i casi.

Recentemente, al fine di monitorare il corretto sviluppo fetale, hanno suscitato notevole interesse le tecniche che sfruttano strumentazioni tridimensionali o addirittura quadridimensionali, impiegate in ambito prenatale (ad esempio, volume e peso Rankin et al. [50], fegato Laudy et al. [51], polmone Pohls et al. [52], placenta e sacca di gestazione Blaas et al. [53]). I lavori che fanno uso di acquisizione volumetriche sono stati orientati principalmente all'individuazione di anomalie fisiche e anatomiche [54, 55]. Sono pochi i lavori orientati all'individuazione della sagittale mediana e ancor meno quelli che riguardano la stima del volume della translucenza nucale.

In Paul et al. [56] viene condotto uno dei primi studi comparativi tra le misurazioni effettuate su sezioni sagittali mediane, identificate manualmente tramite strumentazione bidimensionali e tridimensionali. Lo studio considera 40 donne nel periodo di gestazione compreso tra l'undicesima e la tredicesima settimana. Le misurazioni effettuate sui piani mediani individuati nelle acquisizioni volumetriche sono in accordo con le misurazioni bidimensionali a meno di una differenza compresa tra -0.40mm e 0.60mm. Nel caso in cui il piano sagittale scelto tra quelli acquisiti in modalità volumetrica non sia perfettamente mediano, lo spessore è sotto-stimato con una differenza compresa tra -0.38mm e 0.37mm rispetto alla sezione sagittale mediana tridimensionale. Vengono inoltre riportate le differenze delle misure effettuate su immagini acquisite da un reslicing dei volumi acquisiti e si osservano maggiori errori rispetto al caso bidimensionale. Gli autori riportano alcune considerazioni sulle acquisizioni tridimensionali, sottolineando la difficoltà di acquisizione dei volumi per una corretta misurazione della translucenza nucale che deve essere ben visibile e distinguibile dalla sacca amniotica.

Wee et al. [57] propongono lo strumento di supporto Visualization ToolKit per la visualizzazione grafica interattiva tridimensionale. La metodologia permette la visualizzazione interna delle strutture con la possibilità di selezionare il piano di taglio per la ricerca dei marker ecografici e la loro misurazione come nel caso della translucenza nucale, della quale

ne propongono una valutazione grafica qualitativa.

Un altro studio comparativo tra le misurazioni effettuate con acquisizioni classiche bidimensionali rispetto a quelle di nuova generazione è riportato in Cho et al. [58]; le misurazioni dello spessore della translucenza nelle acquisizioni tridimensionali sono eseguite con il software VolumeNT. Il dataset è composto da 130 acquisizioni, ma soltanto in 93 casi è stato possibile eseguire la misura della translucenza con il software VolumeNT a causa della posizione fetale, dei movimenti fetali durante l'acquisizione e della non chiara visibilità della translucenza; nel 95% e 90% dei casi le misurazioni sono state possibili con i convenzionali metodi bidimensionali e tridimensionali, rispettivamente. Nei casi bidimensionali sono stati rispettati i requisiti descritti in [1] mentre per il caso tridimensionali la sezione sagittale mediana è stata ottenuta ricostruendola dai piani acquisiti tramite la modalità multiplanare. I risultati mostrano che la differenza tra le misurazioni bidimensionali e VolumeNT è compresa in un intervallo tra -0.9mm e 1.0mm, mentre nel caso bidimensionale e tridimensionale è compresa in un intervallo tra -0.8mm e 0.8mm. Gli autori concludono che il programma VolumeNT è in accordo con le misure bidimensionali e tridimensionali e rappresenta un strumento di supporto per il personale medico sebbene il migliore strumento per esaminare la translucenza nucale sia l'acquisizione bidimensionale.

Nel lavoro presentato da Chena et al. [59] si propone una metodologia gerarchica per l'identificazione automatica della sezione sagittale mediana. Sotto l'ipotesi che la testa del feto sia approssimabile ad una sfera, sono calcolate le distribuzioni di probabilità necessarie e sono identificati i punti di interesse attraverso la trasformata di Hough combinata ad un approccio mean shift per stabilire il numero di elementi necessari nella matrice di accumulazione. La metodologia restituisce una accuratezza del 74.00% su un dataset di 1000 immagini sintetiche e del 57.68% su un dataset di 241 immagini reali.

Nel recente lavoro di Cho et al. [60] è stato condotto uno studio comparativo sull'effetto di alcuni parametri, come tissue harmonic imaging (THI), speckle reduction filter (SRF) e sulla misurazione della translucenza in dati acquisiti con strumentazioni tridimensionali. Sono state condotte tre tipologie di misurazioni legate all'ingrandimento dell'immagine: (caso A) quando la testa e il torace del feto occupano meno del 50% dell'immagine, (caso B) quando l'intero feto è contenuto nell'immagine, (caso C) quando la testa e il torace occupano più del 75% dell'immagine. I risultati sono stati ottenuti su 79 soggetti e come ground truth sono

state considerate le misurazioni eseguite su acquisizioni bidimensionali secondo i requisiti della FMF. Sono riportati risultati delle misurazioni eseguite su piani distanti da quello sagittale mediano e si evince che per scostamenti superiori a 40° lo spessore della translucenza è sottostimato. I risultati riportati per le tre tipologie di ingrandimento mostrano una differenza rispetto alla ground truth compresa negli intervalli $(-0.5\text{mm}, 0.3\text{mm})$, $(-0.3\text{mm}, 0.3\text{mm})$ e $(-0.2\text{mm}, 0.2\text{mm})$ rispettivamente per (A), (B) e (C). Ulteriori risultati sono riportati per diverse combinazioni sull'uso di VolumeNT, THI e SRF.

Nonostante vi siano delle correlazioni tra le misure ottenute da acquisizioni in modalità bidimensionale rispetto alla modalità tridimensionale, non sono stati ancora ottenuti risultati che permettano di sostituire le modalità di acquisizione di sezioni sagittali mediane con strumentazioni bidimensionali con quelle individuate in acquisizioni volumetriche. Inoltre le alte frequenze con cui si opera in modalità tridimensionale e quadridimensionale hanno una bassa potenza di penetrazione e i movimenti fetali, che influiscono negativamente sulla qualità dell'acquisizione, insieme agli alti costi delle strumentazioni rallentano la diffusione di queste tecnologie.

3.2 Applicativo

In questo capitolo sarà descritta la metodologia specifica per l'analisi di sequenze video, acquisite tramite ecografo durante il primo trimestre di gravidanza e finalizzata all'individuazione della sezione sagittale mediana, della NT e della sua misurazione. Le sequenze video analizzate, senza perdita di generalità, sono contestualizzate al problema della misurazione della translucenza e l'acquisizione è stata effettuata nel periodo di gestazione che ricade tra l'undicesima e la tredicesima settimana di gravidanza, nel rispetto del protocollo definito dalla FMF.

La sezione sagittale mediana (Figura 3.1), descritta nel capitolo I, costituisce uno dei requisiti fondamentali previsti dal protocollo FMF e richiede particolare attenzione poiché una misurazione su una sezione sagittale non mediana potrebbe portare ad una misurazione falsata con una conseguente diagnosi non corretta. Ad oggi, l'individuazione della sezione sagittale mediana è affidata alla competenza dell'esperto; si evince pertanto una forte soggettività nell'individuazione dell'immagine corretta da analizzare.



Figura 3.1: Sezione sagittale mediana, il rettangolo evidenzia la translucenza nucale.

La metodologia è composta dai seguenti tre moduli principali:

- individuazione della sezione sagittale mediana e relativa validazione;
- individuazione della translucenza nucale e relativa validazione;
- misurazione della translucenza nucale e relativa validazione.

Una varietà di approcci con differenti finalità è descritta in letteratura: dalla soppressione del rumore con tecniche orientate all'analisi di basso livello, alla diagnostica semi-automatica e con metodi che contemplano tecniche di *machine learning*. Un aspetto non semplice da affrontare è la segmentazione che permette di isolare le regioni di interesse dalle quali è possibile estrarre le informazioni necessarie per una efficiente analisi dell'immagine.

3.2.1 Individuazione della sezione sagittale mediana

Data una sequenza video, l'obiettivo di questa prima parte della metodologia consiste nell'identificare, se presenti, i fotogrammi che rappresentano una sezione sagittale mediana. Ricordiamo che il piano sagittale è un piano che attraversa in senso antero-posteriore e divide un corpo in due metà simmetriche, sinistra e destra. La sezione sagittale mediana rappresenta la proiezione dei tessuti e degli organi che il piano attraversa. Nel nostro caso, si considera solo una parte della sezione sagittale per la misurazione della translucenza e pertanto saranno

rappresentate le proiezioni del cranio, la regione nucale, la parte superiore del torace con eventuali tessuti. Nello specifico si acquisisce una sezione sagittale mediana quando la sonda ecografica è posizionata in corrispondenza della falce cerebrale che divide il plesso corioideo in due metà simmetriche. Qualora la sonda non fosse posizionata correttamente si otterrebbe una sezione sagittale, ma non mediana (Figura 3.2).

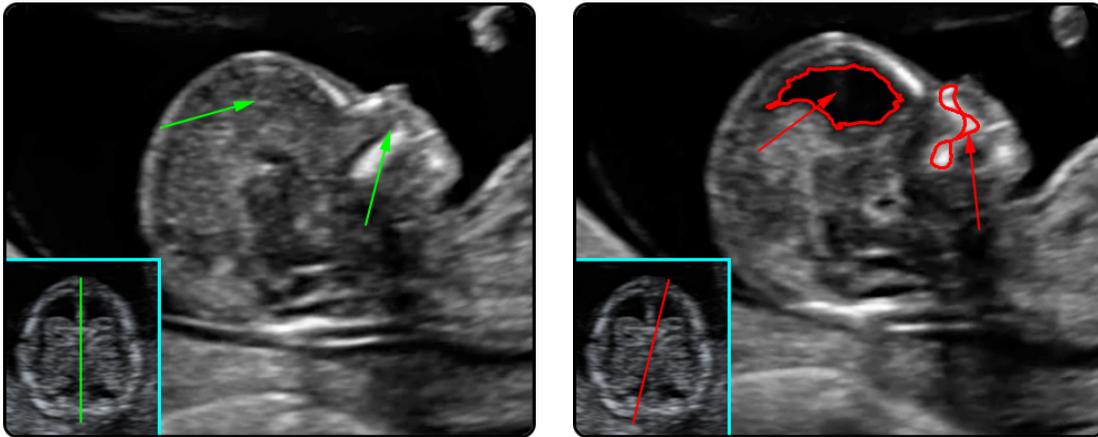


Figura 3.2: Sezione sagittale mediana, le frecce indicano la regione ecogena del plesso e l'assenza del ramo verticale nell'osso mascellare (sinistra). Sezione sagittale non mediana: le frecce indicano la regione anecogena del plesso e il ramo verticale dell'osso mascellare (destra).

Gli elementi che permettono di stabilire se una sezione sagittale è anche mediana sono dipendenti dall'osso mascellare e dal plesso corioideo. L'individuazione e l'analisi delle caratteristiche di queste regioni di interesse sono i passi fondamentali del primo modulo. Ecograficamente, una sezione sagittale non-mediana si distingue da una mediana per la presenza del ramo verticale dell'osso mascellare e del plesso corioideo; è sufficiente che solo uno di questi due elementi sia presente nell'immagine affinché la sezione sia definita non mediana. Come mostrato nella Figura 3.2, laddove è presente il plesso corioideo sarà visibile una regione anecogena a ridosso dell'osso frontale, mentre per l'osso mascellare la presenza del ramo verticale si manifesterà con un tratto iperecogeno, che si estende quasi ortogonalmente al cosiddetto "processo frontale" (Figura 1.27). In Figura 3.3 è riportato il diagramma di flusso che partendo da una sequenza video individua, se presenti, i fotogrammi rappresentanti una sezione sagittale mediana.



Figura 3.3: Diagramma dei moduli funzionali della metodologia proposta.

Analisi del profilo fetale

L'analisi del profilo fetale è il primo e unico modulo che prevede l'elaborazione dell'intera sequenza video per determinare quale sia il profilo fetale, ovvero se l'acquisizione avviene dal lato sinistro o destro (Figura 3.4). Poiché i restanti moduli funzionali effettueranno l'analisi sui singoli fotogrammi in maniera indipendente, dal punto di vista tecnologico è possibile adottare un approccio parallelo al fine di migliorare le prestazioni in termini di tempo. Tale modulo deve considerarsi preliminare perché le informazioni estratte non sono effettivamente utili al fine della misurazione, ma permetteranno di limitare l'analisi a specifiche regioni.

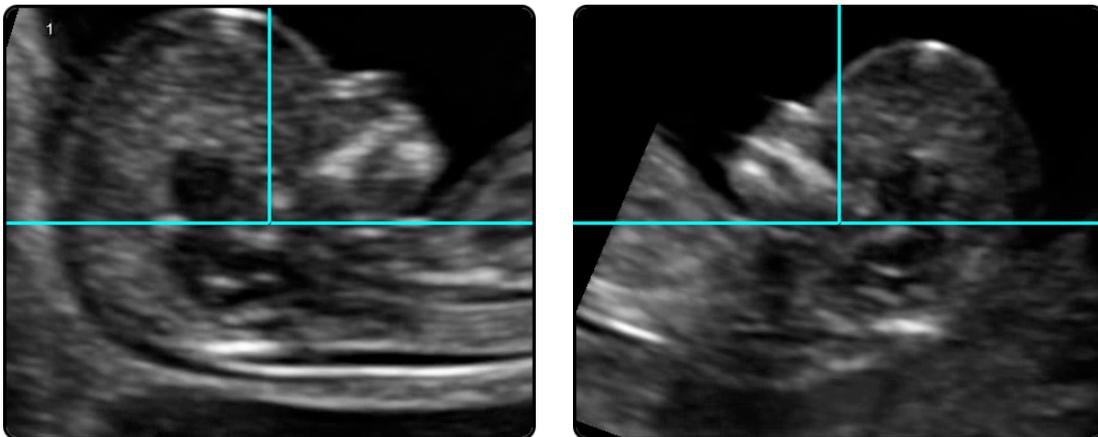


Figura 3.4: Profilo fetale destro (sinistra). Profilo fetale sinistro (destra).

La Figura 3.4 evidenzia che i tessuti ossei presentano un'alta impedenza alle onde acustiche e sono rappresentate ecograficamente da regioni più luminose (iperecogene); in particolare si riscontra una alta percentuale di componenti iperecogene nella parte anteriore del cranio (su

cui insistono l'osso nasale, l'osso mascellare, il mento e l'osso frontale). Sarà dunque identificato il quadrante superiore in cui sono presenti il maggior numero di componenti iperecogene e quindi stabilito il verso del profilo. Non tutti i fotogrammi che compongono la sequenza video rappresentano il profilo fetale e durante l'acquisizione l'esperto può casualmente compiere movimenti bruschi e irregolari che si aggiungono a quelli fisiologici del feto, alterando così la corretta acquisizione del profilo (Figura 3.5). Per questo motivo si è resa necessaria un'analisi globale del video per determinare il “verso” del profilo: abbiamo sperimentalmente verificato che è necessario identificare il quadrante superiore in cui è presente il maggior numero di componenti iperecogene.



Figura 3.5: Assenza del profilo fetale.

Tali componenti iperecogene sono individuate con l'applicazione di una particolare trasformata wavelet, nota come “à trous” che costituisce non solo un metodo semplice e veloce per effettuare la segmentazione, ma consente pure di superare il problema della riduzione del rumore mantenendo nel contempo la risoluzione. Richiamando alcuni concetti esposti nel capitolo II e supponendo che l'immagine in input sia esprimibile come il prodotto scalare delle intensità del livello di grigio g e la funzione di scalatura ϕ :

$$I_0(\mathbf{p}) = I(\mathbf{p}) = \langle g(\mathbf{q}), \phi(\mathbf{q} - \mathbf{p}) \rangle \quad (3.2.1)$$

dove l'iterazione i -esima dell'analisi è espressa da:

$$\frac{1}{4}\phi\left(\frac{\mathbf{p}}{2^i}\right) = \sum_{\mathbf{q} \in \mathbb{D}_\infty} \ell_{\mathbf{q}} \phi\left(\frac{\mathbf{p}}{2^{i-1}} - \mathbf{q}\right) \quad (3.2.2)$$

in cui il pixel \mathbf{p} spazzola l'intera immagine. Nella nostra implementazione, applicheremo il seguente kernel come filtro passa basso ℓ che equivale a scegliere come funzione di scalatura la funzione B_s -spline con $s = 1$:

$$\ell = \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} = \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

Le approssimazioni dell'immagine I sono espresse dalla seguente relazione:

$$I_i(\mathbf{p}) = \frac{1}{4^i} \left\langle f(\mathbf{q}), \phi\left(\frac{\mathbf{q} - \mathbf{p}}{2^i}\right) \right\rangle = \sum_{\mathbf{q} \in \mathbb{D}_\infty} \ell_{\mathbf{q}} I_{i-1}(\mathbf{p} + 2^{i-1} \mathbf{q}). \quad (3.2.3)$$

L' i -esimo piano wavelet sarà calcolato come la differenza tra l'approssimazione I_{i-1} e I_i :

$$W_i(\mathbf{p}) = I_{i-1}(\mathbf{p}) - I_i(\mathbf{p})$$

Al variare dei piani i , il kernel sarà ingrandito in modo da contenere $(2^i + 1) \times (2^i + 1)$ elementi, pur mantenendo costante il numero di elementi non nulli. Ciò equivale ad applicare direttamente ad I_0 un opportuno kernel K di dimensione $(2^i - 1) \times (2^i - 1)$. Tale trasformata permette di evidenziare nell'immagine strutture di determinate dimensioni in funzione di i : piccoli oggetti saranno distinguibili nei primi piani, mentre all'aumentare di i saranno esaltate le componenti con dimensioni maggiori (Figura 3.6).

Limitatamente alla risoluzione delle immagini in nostro possesso, abbiamo verificato sperimentalmente che le componenti relative ai tessuti ossei del cranio sono ben identificate dal piano W_6 . Inoltre va osservato che poiché piani attigui poco si differenziano tra loro, informazione simili si trovano nei piani W_5 e W_7 (Figura 3.6). La segmentazione in tali immagini è eseguita tramite semplice sogliatura sul piano W_6 :

$$C(\mathbf{p}) = \{\mathbf{p} : W_6(\mathbf{p}) \geq \mu(W_6) + 2\sigma(W_6)\}$$

dove $\mu(W_6)$ e $\sigma(W_6)$ indicano rispettivamente la media e la deviazione standard dei livelli di luminosità del piano W_6 (Figura 3.7).

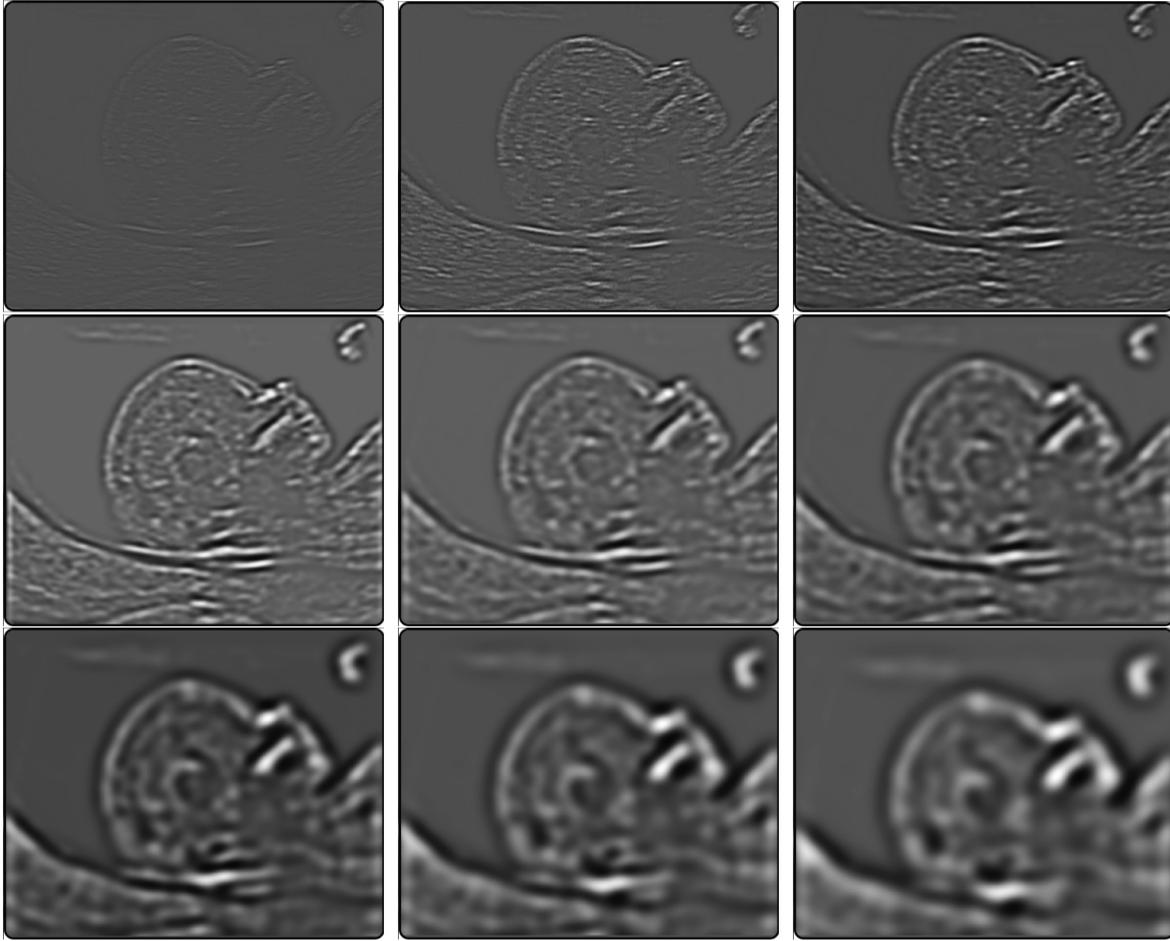


Figura 3.6: Piani wavelet per valori di $i = \{1, \dots, 9\}$ da sinistra verso destra dall'alto verso il basso.

Nel seguito considereremo le componenti che ricadono nella metà superiore dell'immagine e dividendo idealmente l'immagine verticalmente in due metà, valuteremo per maggioranza su tutti i fotogrammi, quale è il quadrante che presenta il maggior numero di componenti (Figura 3.8). Una volta identificato tale quadrante è banale stabilire quale sia il profilo del feto. Senza perdita di generalità, descriveremo la metodologia considerando il solo profilo destro. Analoghe considerazioni valgono anche per il caso simmetrico del profilo sinistro.

Le pareti dell'utero, rappresentato ecograficamente da una regione ecogena (Figura 3.9-*sinistra*), fanno parte di quei tessuti materni che è possibile individuare nelle sezioni sagittali. Essendo adiacente al liquido amniotico presenta un brusco cambio di frequenza ed è dunque intercettato come componenti strette ed allungate nei piani wavelet segmentati (Figura



Figura 3.7: Piano wavelet segmentato profilo destro (sinistra). Piano wavelet segmentato sinistro (destra).

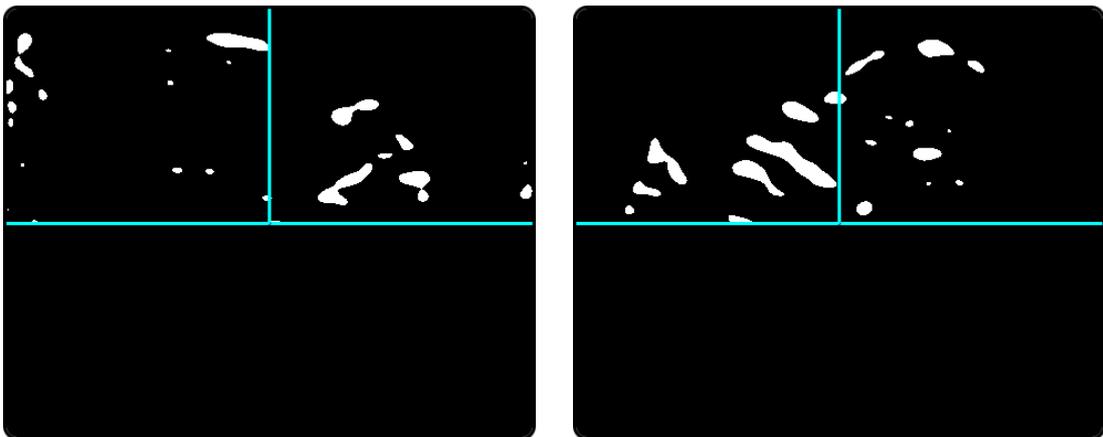


Figura 3.8: Componenti quadranti per il profilo destro (sinistra) e sinistro (destra).

ra 3.9-*destra*). Sarà questa caratteristica, insieme a ulteriori relazioni spaziali, a permettere di individuare l'area di interesse, mentre la successiva fase analizzerà soltanto la sua zona sottostante (Figura 3.10). Questo modulo funzionale è stato sviluppato per una corretta individuazione delle regioni di interesse, come nel caso del plesso corioideo che può essere identificato come una regione con elevata simmetria radiale.

Identificazione dell'osso mascellare

La sezione della metodologia destinata ad analizzare l'osso mascellare ha lo scopo di verificare la presenza del ramo verticale che implica l'esclusione del fotogramma nelle successive fasi

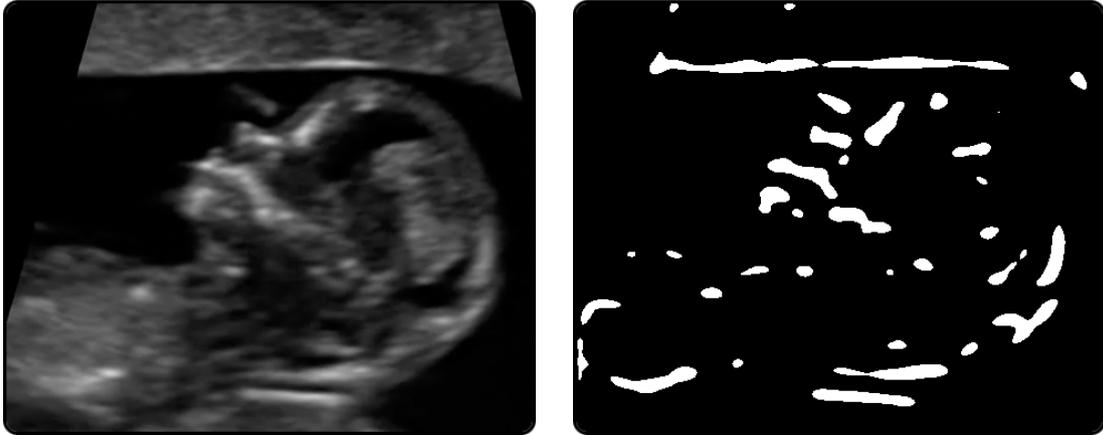


Figura 3.9: Ecografia fetale con tessuti materni visibili in alto (sinistra) e relativo piano wavelet segmentato (destra).



Figura 3.10: Ecografia fetale con soppressione dei tessuti materni.

della metodologia: è sufficiente la presenza di uno solo tra ramo verticale e plesso corioideo affinché la sezione sia definita non mediana, evitando la successiva valutazione.

A tal fine è stata addestrata una rete neurale multistrato feed forward (per maggiori dettagli sull'addestramento si rimanda al capitolo 4) con l'obiettivo di apprendere le caratteristiche che permettono di discriminare tra le classi rappresentative del ramo verticale visibile o non visibile (Figura 3.11). La rete neurale riceve in ingresso l'immagine binaria e ogni neurone del livello di output restituirà la probabilità di appartenenza della componente alla classe che esso rappresenta; la componente sarà associata alla classe di appartenenza per cui il valore

di probabilità è maggiore.

Le immagini sono state ottenute con lo stesso procedimento di segmentazione descritto precedentemente per l'analisi del profilo. Inoltre, prima di essere sottoposte alla rete neurale, le immagini binarie sono processate al fine di sopprimere tutte le componenti che ricadono al di fuori del quadrante.

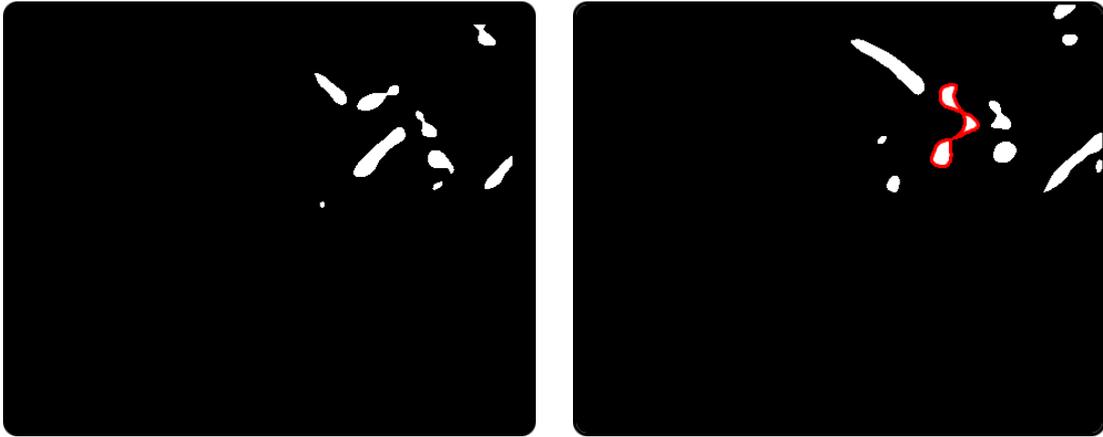


Figura 3.11: Immagine input appartenente alla classe C1, ramo non visibile (sinistra). Immagine input appartenente alla classe C2, ramo visibile (destra).

Analisi dell'osso mascellare

Le immagini che sono state identificate come appartenenti alla classe C1 (ovvero ramo verticale non visibile) sono poi processate da una seconda rete neurale per classificare le varie componenti. Le immagini in ingresso sono le stesse processate dalla prima rete neurale con l'ulteriore vincolo di appartenenza alla classe C1.

La seconda rete neurale è stata addestrata in modo da distinguere quattro classi: naso, mento, mandibola e "altro" (in quest'ultima classe ricadono tutte quelle componenti che non appartengono alle precedenti), come mostrato in Figura 3.12.

Questo modulo sarà di supporto all'individuazione dell'osso frontale: identificando la componente "mandibola" sarà possibile limitare ulteriormente la ricerca per l'analisi del plesso.



Figura 3.12: Immagine input seconda rete neurale: giallo: osso nasale, azzurro: osso mascellare, fucsia: mento, bianco: “altro”.

Individuazione della regione frontale

Quando la falce cerebrale è pienamente intercettata dal fascio di onde acustiche, sarà visualizzata ecograficamente una regione maggiormente ecogena (non è un tessuto osseo) estesa fino all'osso frontale (Figura 3.2-sinistra). La regione anecogena invece è visibile quando il fascio acustico intercetta il liquido cerebrale nella regione tra la falce e il plesso (sinistro o destro). Poiché tale regione è posizionata nella parte superiore del cranio e quest'ultima è approssimabile con una semi-circonferenza, sorge l'esigenza di individuare i centri di simmetria radiale. La variabilità delle dimensioni del cranio fetale, nel periodo gestazionale compreso tra l'undicesima e la tredicesima settimana è limitata, consentendo di applicare la cosiddetta “radial symmetry transform” [27]. L'algoritmo originale è stato modificato in modo da pesare maggiormente gli archi di circonferenza compresi tra 60° 120° corrispondenti alla regione centrale del cranio che comprende in parte l'osso frontale e in parte quello parietale (Figura 3.13). Le coordinate del punto con grado massimo di simmetria, limitatamente alla porzione delimitata dalle semirette uscenti dal vertice inferiore destro del minimo rettangolo di ricoprimento dell'osso mascellare, restituiscono pure il raggio della circonferenza corrispondente (Figura 3.13).

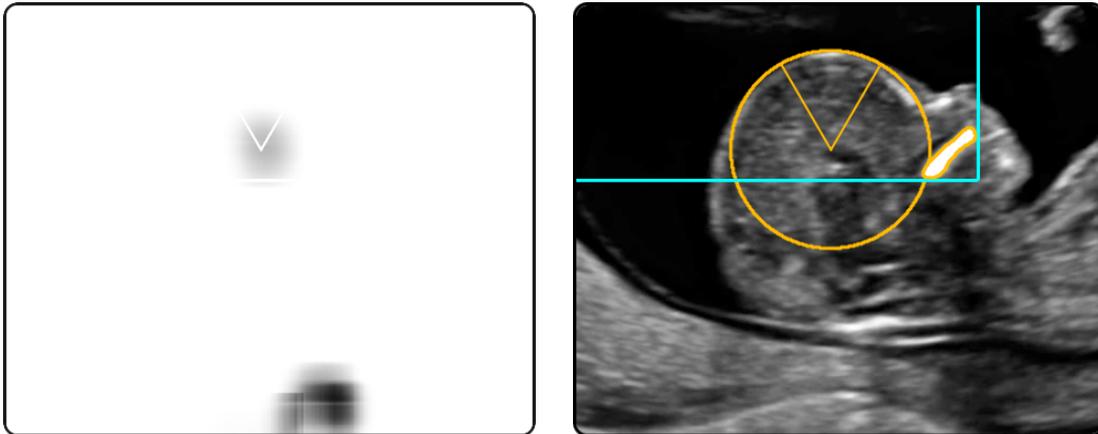


Figura 3.13: Output della trasformata simmetria (sinistra). In evidenza la regione con alta simmetria radiale delimitata dalla circonferenza (destra).

Analisi della regione frontale

Questo modulo funzionale determina infine se nella sezione sagittale corrente non è visibile il plesso corioideo, affinché possa essere definita mediana. L'analisi dell'ecogenicità della regione frontale è stata ottenuta con lo sviluppo di una procedura in cui dapprima è delimitata la regione di interesse e successivamente è valutata l'intensità dei pixel al suo interno. Tale regione corrisponde al settore circolare delimitato dalle circonferenze con raggi 25 e 85 e centro nella circonferenza precedentemente individuata, compreso tra 10° - 100° nel caso in cui l'immagine raffigura un profilo destro (simmetricamente 80° - 170° per il profilo sinistro) come in Figura 3.14.

I pixel appartenenti al settore circolare sono etichettati in accordo alle funzioni di distribuzione di probabilità stimate e descritte nel prossimo capitolo 4: R (anecogene), G (ecogene) B (incerto). L'ulteriore classe considerata B (incertezza) è necessaria per l'alta variabilità in luminosità delle sequenze video. Assumiamo l'assenza del plesso corioideo se il numero dei pixel ecogeni e anecogeni soddisfano il seguente test:

$$\begin{cases} \frac{\#anechogenic}{\#overall} \leq \phi \\ \frac{\#echogenic}{\#overall} \geq \varphi \end{cases}$$

dove $\#overall$ si riferisce al numero di pixel "incerti" all'interno del settore circolare, con

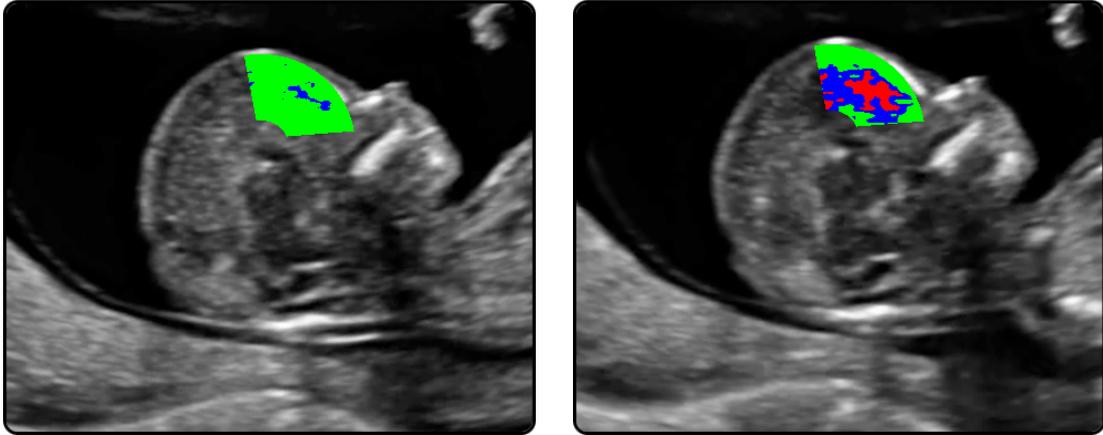


Figura 3.14: Analisi della regione frontale, in rosso i pixel anecogeni, in verde i pixel ecogeni, in blu i pixel “incerti”. La regione non è anecogena (sinistra). La regione è anecogena (destra).

$\varphi = 0.11$ e $\phi = 0.43$ valori empirici (vedi Capitolo 4 per la descrizione degli esperimenti). Superato tale test, la sezione sagittale può essere definita mediana.

3.2.2 Individuazione della translucenza nucale

Tenendo conto delle informazioni spaziali ricavate dai moduli precedenti è finalmente individuata la regione di interesse contenente la translucenza nucale. Supportati dalle relazioni anatomiche e morfologiche del cranio, dalle ossa in esso racchiuse (mandibola) e dalla translucenza nucale, limitiamo la ricerca della regione di interesse considerando il centro e il raggio della circonferenza individuata nel modulo “Individuazione della regione frontale” e le coordinate dell’osso mascellare individuato nel modulo “Analisi dell’osso mascellare”, come mostrato in Figura 3.16-*sinistra*. Prima di procedere nella descrizione dei dettagli tecnici, dobbiamo fare una considerazione riguardo le componenti ecogene che delimitano la translucenza nucale. Come in Figura 3.1, la componente superiore è sempre facilmente individuabile poiché la sua ecogenicità è solitamente la maggiore. Purtroppo, la componente inferiore non sempre è facilmente individuabile. Invece, la translucenza nucale compresa tra queste due regioni risulta sempre anecogena anche se va ricordato che nelle sue prossimità possono essere presenti altre strutture fortemente anecogene (ad esempio, nel caso della cisterna magna). Si è reso quindi necessario modificare il valore di soglia secondo la seguente regola (Figura 3.15):

$$C(\mathbf{p}) = \{\mathbf{p} : W_6(\mathbf{p}) \leq \mu(W_6) + 2\sigma(W_6)\} \quad (3.2.4)$$



Figura 3.15: Negazione del piano wavelet già binarizzato.

Tra le componenti che ricadono nella regione messa in evidenza nella Figura 3.16-*sinistra*, la translucenza nucleare è associata alla componente con area massima. La regione di interesse è migliorata in base all'area occupata dalla componente scelta, come mostrato in Figura 3.16-*destra*.

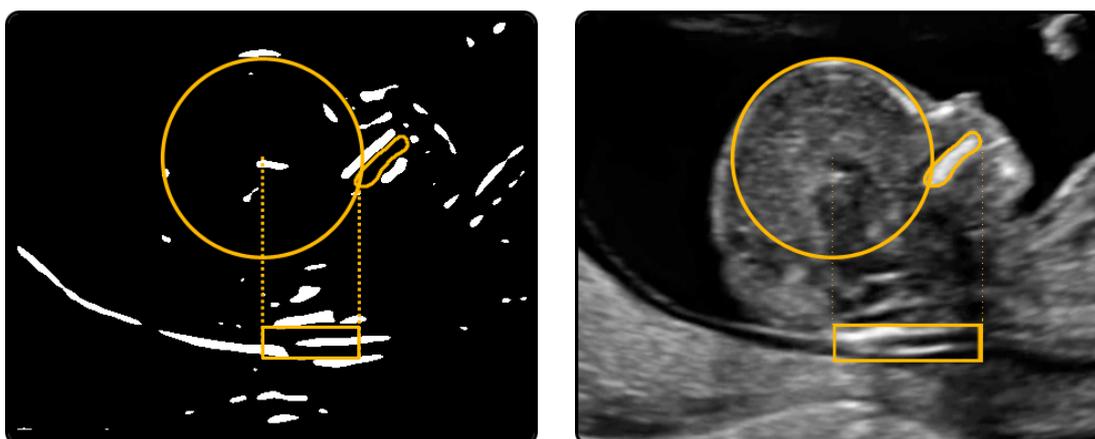


Figura 3.16: Individuazione della regione nucleare: le componenti anecogene interessate (*sinistra*) e la regione nucleare finale messa in evidenza (*destra*).

3.2.3 Misurazione della traslucenza nucale

Generalmente le strutture nelle immagini ecografiche non hanno un contorno ben definito e nel caso specifico della traslucenza nucale, il cambiamento di luminosità gioca un ruolo fondamentale a causa di numerosi e repentini cambi di luminosità (per esempio, dovuti allo speckle noise). Il metodo sviluppato è basato sull'utilizzo di informazioni estratte tramite trasformata wavelet a scale diverse. L'immagine di partenza è preliminarmente decomposta in bande di frequenza, secondo i paradigmi della analisi multirisoluzione. Le informazioni sui bordi vengono dunque propagate tra le varie scale: i bordi alle alte frequenze potrebbero apparire segmentati ed essere invece contigui nelle basse frequenze.

Singolarmente le maschere binarie non riescono a garantire un buon compromesso tra l'estensione delle componenti e la definizione del bordo a causa dell'elevata variabilità (in luminosità, ma anche in rumore) delle sequenze video ecografiche. Pertanto le maschere binarie ottenute da piani consecutivi, sono unite insiemisticamente in una nuova maschera binaria che comprende le informazioni presenti nelle diverse scale. Tale tecnica preserva le dimensioni delle componenti e mantiene contemporaneamente un livello sufficiente di dettaglio presente nelle alte frequenze. L'intera procedura risulta robusta anche in presenza di rumore di tipo speckle, garantendo componenti con bordi continui e lineari. In particolare, sono scelti i piani W_4 , W_5 e W_6 (Figura 3.6) e sogliaati secondo (3.2.5) (con $k = 1.2$, individuato sperimentalmente e per i dettagli si rimanda al Capitolo IV), ottenendo le maschere binarie B_4 , B_5 e B_6 (Figura 3.17).

$$B_i(\mathbf{p}) = \{\mathbf{p} : W_i(\mathbf{p}) \leq \mu(W_i) + k\sigma(W_i)\} \quad (3.2.5)$$

Le maschere C_i sono quindi combinate insieme mediante votazione per ciascun pixel di almeno due su tre, per restituire una nuova maschera contenente le strutture principali (Figura 3.17-*in basso a destra*).

$$B(\mathbf{p}) = \{\mathbf{p} : \sum_{i=4}^6 B_i(\mathbf{p}) \geq 2\} \quad (3.2.6)$$

La traslucenza nucale è quindi facilmente identificabile (Figura 3.18); l'applicazione di operatori morfologici classici come hit or miss nella maschera binaria finale ne estraggono agevolmente il bordo (Figura 3.19). Il *diametro* è definito come la distanza tra i pixel ap-

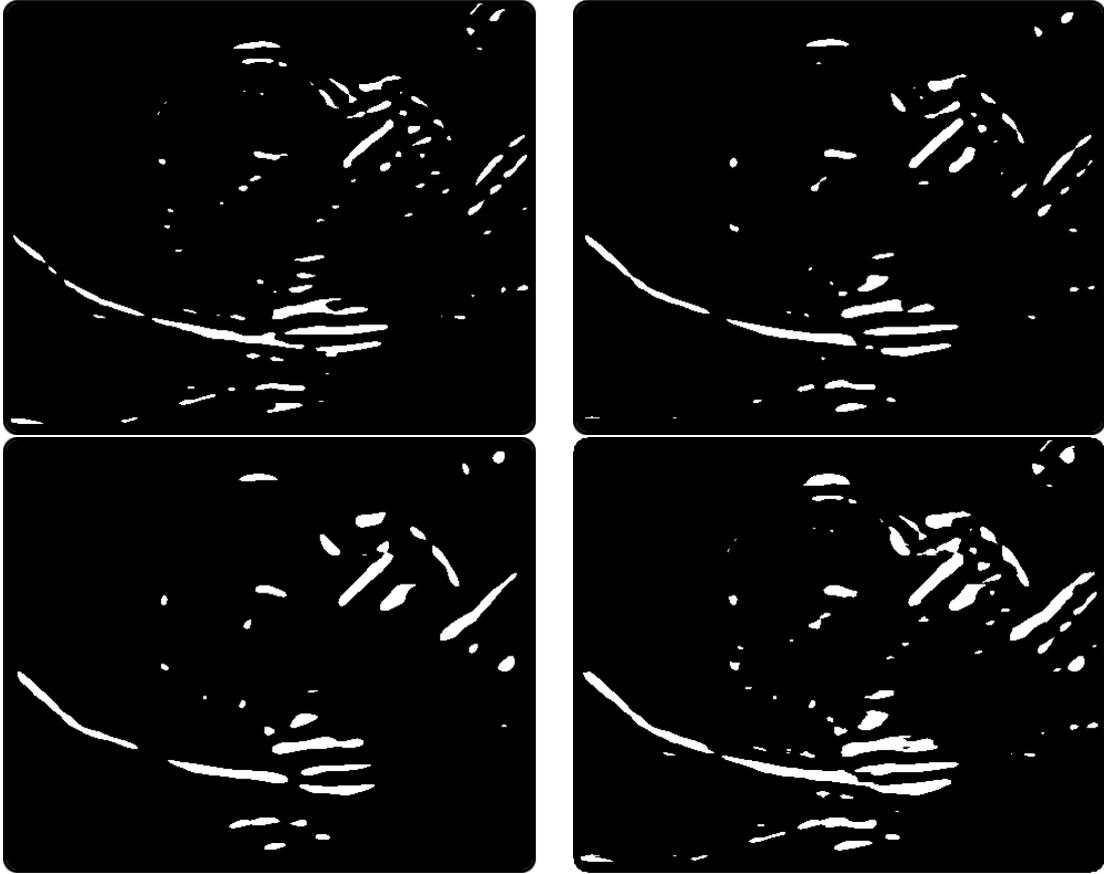


Figura 3.17: I piani wavelet W_4 , W_5 e W_6 e la loro combinazione tramite votazione.

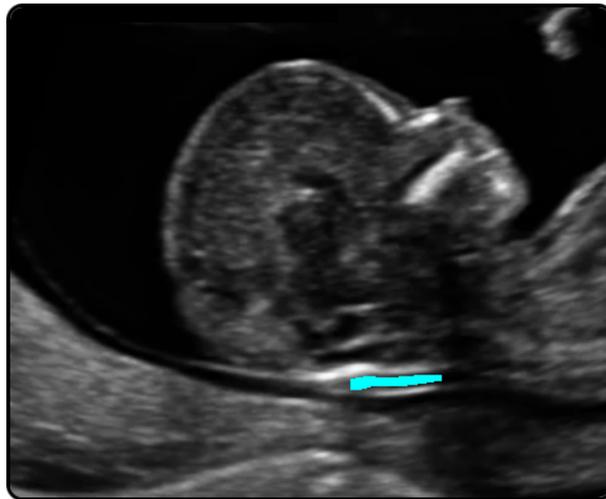


Figura 3.18: La componente identificata come traslucenza nucale in evidenza (azzurro).

partenenti ai bordi inferiore e superiore; il diametro massimo, secondo il protocollo FMF, corrisponde alla misura da annotare per il fotogramma in analisi (Figura 3.19).



Figura 3.19: Estrazione del bordo, in rosso, che delimita la translucenza nucale e dei diametri rappresentativi (segmenti verticali), in verde il diametro massimo.

Capitolo 4

Risultati

4.1 Descrizione dei dati in analisi

In questo capitolo saranno esposti i risultati ottenuti da una estesa sperimentazione su sequenze video acquisite tramite l'ecografo GE Voluson E8. Le sequenze video sono state acquisite dal "Centro Diagnosi Prenatale" di Palermo; la validazione dei risultati ottenuti è stata curata da medici esperti accreditati dalla Fetal Medicine Foundation, i quali periodicamente si sottopongono a test di valutazione. Le acquisizioni sono state condotte, nel rispetto dell'anonimato, su donne in gravidanza il cui periodo di gestazione è compreso tra l'undicesima e la tredicesima settimana per la misurazione della traslucenza nucale. In particolare è stato sviluppato un ambiente software, di supporto alla diagnosi prenatale, per l'individuazione dei fotogrammi più opportuni per l'identificazione e la misurazione automatica della traslucenza. Lo spessore della traslucenza nucale, aumentato o nella norma rispetto ai parametri di riferimento, non è stato un elemento discriminante, infatti i dati sono stati acquisiti e processati senza nessuna distinzione.

Le sequenze video presentano una ampia variabilità sia nella morfologia dei soggetti in esame, sia nella qualità delle immagini in termini di contrasto e di rumore. La frequenza degli ultrasuoni emessi varia, a seconda della costituzione del soggetto sottoposto all'esame ecografico, in un range compreso tra i 3 e i 10 MHz; si fa presente che la costituzione fisica, in particolare gli strati adiposi addominali delle donne che si sottopongono all'esame ecografico, è un elemento determinante per una buona rappresentazione del feto. Tutte le acquisizioni sono

state registrate in formato digitale di 640×480 pixel con il più basso fattore di compressione consentito dal codec H.264 [61] al fine di ridurre il più possibile artefatti. Sebbene i video siano percepiti su scala di grigio abbiamo estratto il solo canale di luminanza dallo spazio colore YUV ottenendo una scala di grigio nell'intervallo [0 255]. La durata delle acquisizioni, ovvero il numero di fotogrammi per sequenza, è variabile. L'intera metodologia è stata implementata e sviluppata in ambiente Matlab. Buona parte del lavoro descritto nella presente tesi, con particolare riguardo alle varie fasi utili alla determinazione della sezione sagittale mediana, è attualmente sottomessa a rivista [62], mentre ulteriori aspetti della ricerca in tale ambito ha già portato alla pubblicazione del lavoro [48].

4.2 Addestramento reti neurali

Al fine di sviluppare la metodologia esposta nel capitolo III, è stato condotta una valutazione preliminare per testare la capacità di riconoscimento delle due reti neurali multistrato feed forward: la prima rete è addestrata per la classificazione relativa all'osso mascellare, considerando le classi "ramo verticale non visibile" (C1) e "ramo verticale visibile" (C2); la seconda rete neurale, addestrata limitatamente ai fotogrammi in cui non è presente il ramo verticale, prende in considerazione le classi "osso nasale" (CN), "mento" (CC), "osso mascellare" (CM) e "altri" (CA) con lo scopo di classificare le principali componenti di nostro interesse presenti nel profilo fetale. In generale, l'addestramento di una rete neurale è caratterizzato dalle due fasi principali di apprendimento e di testing. Per entrambi le reti, l'apprendimento è di tipo supervisionato: l'utente fornisce cioè esempi rappresentativi di ciascuna classe, in modo tale da permettere alla rete di apprendere le caratteristiche.

I dati forniti alle reti neurali utilizzati per l'addestramento sono stati corredati di una etichetta identificativa della classe di appartenenza. Il data set disponibile è stato suddiviso in due sottoinsiemi, *training set* e *test set*, ad intersezione vuota. Il training set è stato utilizzato nella fase di apprendimento, durante la quale la rete apprende le principali caratteristiche, modificando opportunamente i pesi della rete; il test set è utilizzato per verificare le prestazioni della rete e quindi la sua capacità di riconoscimento, quando nuovi dati sono presentati. Le reti neurali accettano in input vettori che rappresentano le caratteristiche dei dati, motivo per cui ciascun fotogramma è stato considerato come vettore delle sue caratteristiche. Le

immagini, prima di essere processate dalle reti, sono state pre-elaborate applicando la trasformata wavelet e successivamente binarizzate come descritto nel capitolo III. Il riquadro di riferimento, che conterrà le principali ossa craniche di nostro interesse, è identificato considerando il verso del profilo fetale. In (Figura 4.1) sono mostrate le immagini rappresentative per le due classi. Le reti neurali sono state strutturate su tre livelli con le seguenti configurazioni:

- input: il livello di ingresso ha un numero di nodi pari al numero di pixel dell'immagine;
- hidden: il numero di nodi del livello nascosto è dipendente dal numero di nodi del layer di input;
- output: il livello di uscita ha un numero di nodi pari al numero di classi considerate.

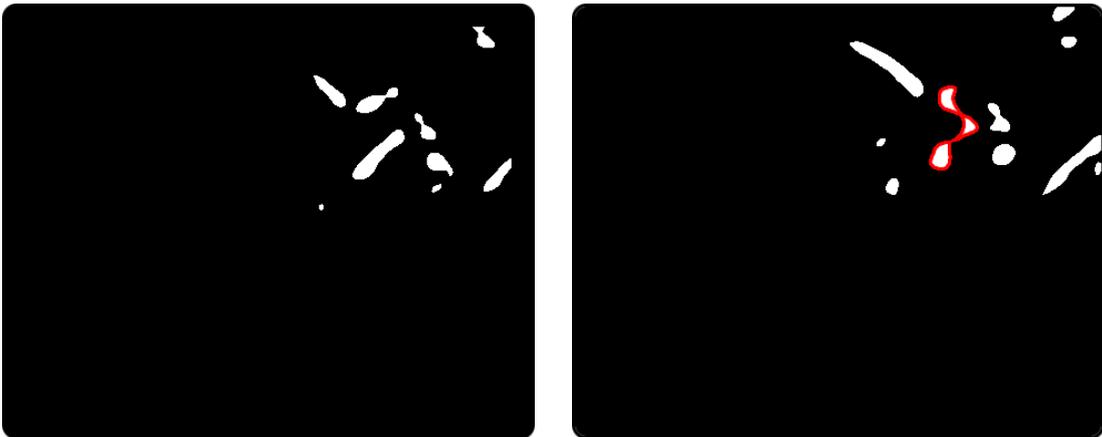


Figura 4.1: Immagine rappresentativa della classe I: ramo verticale non visibile (sinistra). Immagine rappresentativa della classe II: ramo verticale visibile (destra).

Il tipo di apprendimento è basato sull'algoritmo back-propagation e il numero massimo di epoche è stato fissato a 1000; la funzione di errore scelta è la quadratica media (MSE) e affinché possa essere innescato il cosiddetto early stopping, il suo valore minimo è stato posto pari a 1.0^{-6} mentre l'inizializzazione dei pesi è di tipo casuale.

A supporto dell'efficienza della classificazione sono stati condotti dei test di valutazione preliminari sulla capacità di riconoscimento delle reti neurali multistrato feedforward limitatamente al nostro contesto; la matrice di confusione, ottenuta dai test effettuati, per ciascuna rete (Figure. 4.1, 4.2), ha dimostrato che le reti neurali sono in grado di costruire un modello

generalizzato per la classificazione dei dati analizzati. Per validare la capacità delle reti neurali, è stato utilizzato il metodo del cross validation. L'insieme dei dati è stato partizionato in 3 sottoinsiemi disgiunti e ciascun sottoinsieme è utilizzato come training set, mentre i due sottoinsieme restanti sono utilizzati come test set, addestrando così 3 reti neurali per l'intero dataset. Questo procedimento è ripetuto 10 volte, disponendo in modo casuale gli elementi del dataset e quindi dei vari sottoinsiemi formati. Una rappresentazione del procedimento di creazione dei dataset è mostrato in Figura 4.2.

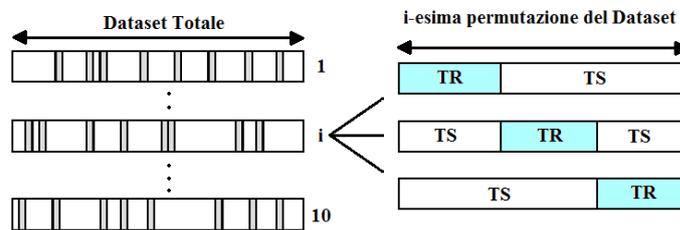


Figura 4.2: Ciascuna delle 10 copie del dataset i cui elementi sono disposti in modo casuale (a sinistra), è suddivisa in 3 partizioni, a turno una sarà il training set (TR) le due restanti il test set (TS) per l'addestramento delle reti (a destra).

Rete Neurale I

L'obiettivo della prima rete neurale è l'apprendimento un modello valido per la classificazione delle due classi: presenza del ramo verticale, assenza del ramo verticale. Per l'addestramento della rete, il dataset preso in considerazione è composto da 831 immagini estratte da 6 filmati, con una equa rappresentazione di entrambe le classi e per ciascuna sequenza video. Riassumendo, ognuna delle 30 coppie (training set, test set) formata dalla disposizione casuale del dataset ha una cardinalità di 831 elementi, di cui 271 appartenenti al training set e 560 al test set, per un totale, sulle 30 reti, di 8130 e 16800 fotogrammi rispettivamente. Di seguito la matrice delle performance cumulativa per le 30 reti neurali (Tabella 4.1).

Rete Neurale II

L'obiettivo della seconda rete neurale è l'apprendimento un modello valido per la classificazione delle quattro classi: osso mascellare, osso nasale, mento e "altri". Per l'addestramento

Tabella 4.1: Matrice delle performance relativa alle 30 reti neurali I.

		Medico	
		si	no
Metodologia	si	86.68%	14.47%
	no	13.32%	85.53%

Addestr. ramo verticale

della rete, il dataset preso in considerazione è composto da 1902 immagini selezionate da 7 filmati, con una equa rappresentazione di ciascuna delle 4 classi per ciascuna sequenza video. Riassumendo, ognuna delle 30 coppie (training set, test set) ha una cardinalità di 1902 elementi, di cui 625 appartenenti al training set e 1277 per il test set, per un totale, sulle 30 reti, di 18750 e 38310 fotogrammi rispettivamente. Nonostante nella metodologia verranno utilizzate le informazioni relative alla sola componente “mandibola”, uno studio preliminare sul numero di classi da considerare ha permesso di stabilire che la capacità di apprendimento delle reti è migliore nel caso in cui sono considerate 4 classi. Di seguito la matrice delle performance cumulativa per le 30 reti neurali (Tabella 4.2).

Tabella 4.2: Matrice delle performance relativa alle 30 reti neurali II.

		Medico			
		mandibola	mento	naso	altro
Metodologia	mandibola	95.89%	1.43%	1.04%	0.65%
	mento	1.68%	96.69%	0.30%	2.96%
	naso	2.05%	0.01%	98.48%	0.30%
	altro	0.38%	1.87%	0.18%	96.09%

Addestramento componenti facciali

4.3 Stima delle funzioni di probabilità

La presenza o assenza del plesso corioideo è determinata analizzando l'ecogenicità della regione di interesse. Per ciascun pixel di tale area è calcolato il grado di appartenenza ad una delle tre possibili classi: anecogeno, iperecogeno e incerta, rappresentate da una funzione di distribuzione di probabilità. Le motivazioni che rendono necessaria l'introduzione della classe incertezza sono legate alla variabilità dei livelli di grigio, in termini di luminosità globale, delle sequenze video. Uno studio preliminare è stato condotto sugli istogrammi dei livelli di grigio associati ai filmati; da tale studio non è emersa alcuna caratteristica discriminante che permetta di classificare le sequenze in base al loro contrasto, da cui i pixel ne ereditano le proprietà; infatti non esiste una soglia adeguata per stabilire se un video presenta un alto o un basso livello contrasto, e di conseguenza non esiste una soglia generica per stabilire se un pixel può essere definito iperecogeno o anecogeno. Infatti la definizione stessa di ecogenicità è limitata al confronto tra due valori di luminosità. Sono state stimate tre funzioni di distribuzione di probabilità rappresentative delle tre classi e tale scelta si è rivelata adeguata in termini di sensibilità, specificità e tempo computazionale. La stima si è basata su un campionamento supervisionato effettuato manualmente su un insieme di immagini, selezionando aree di pixel di dimensione 3×3 (Figura 4.3).

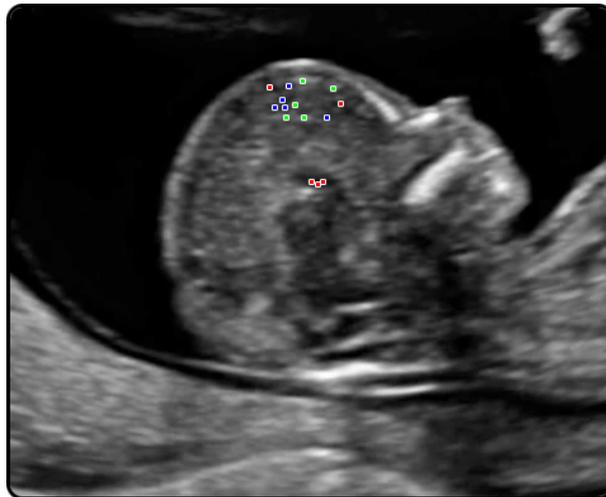


Figura 4.3: Aree campionate: anecogene (rosso), ecogene (verde) and incertezza (blu).

Le immagini, utilizzate per il campionamento, rappresentano sia profili in cui il plesso è

visibile, sia quelli in cui non lo è. Come descritto nella sezione “Analisi della regione frontale”, per ottenere una stima accurata, sono state marcate in ciascuna immagine 5 aree per ciascuna classe, limitatamente alla regione del plesso corioideo presa in considerazione; marcare regioni rappresentanti il tessuto osseo o lo sfondo non sarebbe molto significativo ai fini della stima, in quanto in tale regione non vi sono presenti tessuti con una ecogenicità simile. Calcolata la media e la deviazione standard dei valori di grigio dei pixel selezionati per ciascuna classe, sono state stimate le funzioni di distribuzione di probabilità (Figura 4.4).

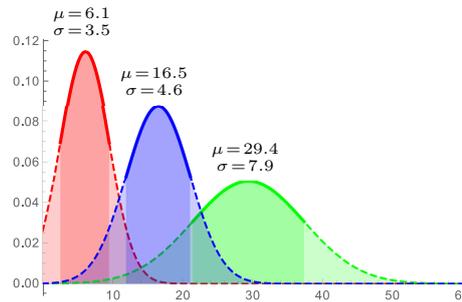


Figura 4.4: Funzioni di distribuzione di probabilità: anecogene (rosso), ecogene (verde) and incertezza (blu) con media μ e deviazione standard δ .

4.4 Stima dei parametri per il test del plesso corioideo

Determinare se il settore circolare identificato nel modulo “Analisi della regione frontale” rappresenta una area anecogena equivale a valutare se il numero di pixel etichettati come ecogeni rappresenta una porzione estesa nell’area considerata. In ultima analisi, il fotogramma è definito sezione sagittale mediana se soddisfa il seguente test:

$$\begin{cases} \frac{\#pixelanecogeni}{\#overall} \leq \phi \\ \frac{\#pixelecogeni}{\#overall} \geq \varphi \end{cases}$$

dove con $\#overall$ indichiamo il numero di pixel totali appartenenti al settore circolare, $\#pixelecogeni$ il numero di pixel etichettati come ecogeni (analogamente per $\#pixelanecogeni$), mentre $\varphi = 0.43$ e $\phi = 0.11$. Inizialmente la stima è stata condotta per un solo valore di soglia ϕ , considerando i valori di threshold da 0.1 a 0.6 per determinare quale fosse il valore

idoneo per tutte le sequenze. I risultati poco incoraggianti raffigurati in Figura 4.5, descrivono un andamento pressocchè casuale nella classificazione; pertanto è stato inserito il secondo parametro φ in modo che anche la classe “incertezza” fosse implicitamente rappresentata. I grafici esprimono l’efficienza del test sul plesso corioideo al variare dei parametri considerati.

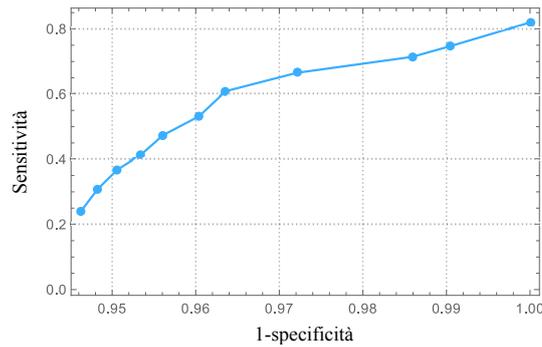


Figura 4.5: Curva ROC al variare del valore di soglia ϕ .

I valori di ϕ e φ sono stati determinati su base “statistica”, facendoli variare in un range [0.05 0.25] per ϕ e in un range [0.40 0.60] per φ (Figura 4.6 e Figura 4.7).

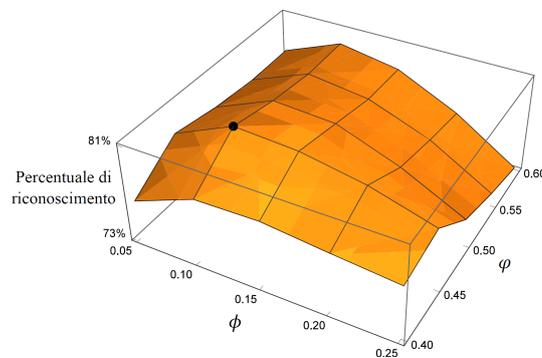


Figura 4.6: Efficienza al variare di ϕ e φ rispettivamente negli intervalli [0.05, 0.25] e [0.40, 0.60] con step 0.05.

Per determinare in modo più preciso ϕ e φ , lo step di variazione è stato ridotto a 0.01, riducendo l’ampiezza degli intervalli di ricerca, in accordo ai valori individuati considerando uno step maggiore (0.05).

Si è verificato che il massimo valore si ha per $\phi = 0.11$ e $\varphi = 0.43$

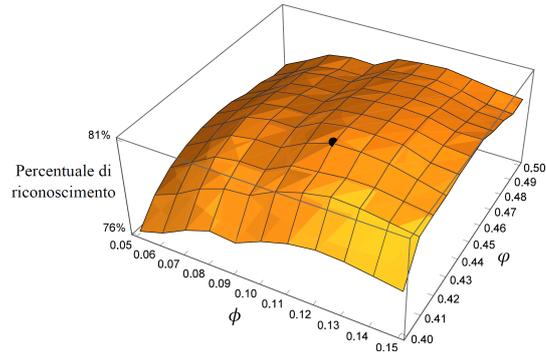


Figura 4.7: Efficienza al variare di ϕ e φ rispettivamente negli intervalli $[0.10, 0.15]$ e $[0.40, 0.50]$ con step 0.01.

4.5 Stima del coefficiente k per l'identificazione della translucenza nucleale

Una analisi quantitativa delle maschere binarie ottenute dai piani wavelet in funzione del parametro k (cfr. (3.2.5)) ha permesso di individuare il miglior valore in modo tale che fosse minimizzato l'errore medio di misurazione automatica rispetto alle misurazioni manuali e contemporaneamente massimizzato il numero di fotogrammi che presentano un errore nel diametro non superiore all'errore medio. Da tale sperimentazione, Figura 4.8, si deduce che il valore più opportuno per il coefficiente moltiplicativo è $k=1.2$ in corrispondenza del quale si ottiene un errore medio pari a 2.28 soddisfatto dal 61.11% delle soluzioni trovate.

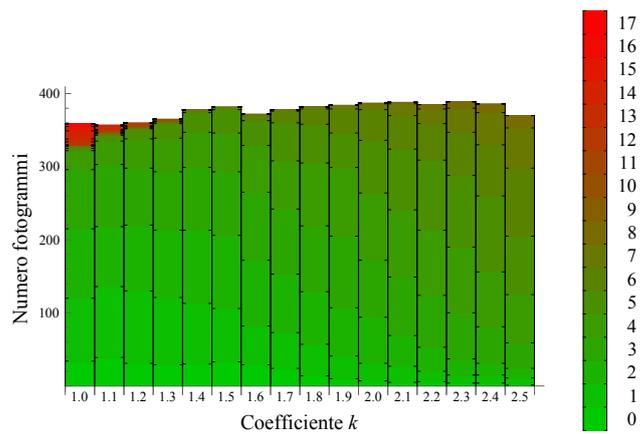


Figura 4.8: Errore commesso dalla metodologia automatica al variare del coefficiente k , la legenda esprime il valore in pixel dell'errore rispettivamente per il colore indicato.

4.6 Risultati

Di seguito verranno mostrati i risultati ottenuti per ciascuna fase della metodologia. Per la prima fase, l'identificazione della sezione sagittale, mostreremo la capacità di riconoscimento all'interno di una sequenza video dei fotogrammi che rappresentano sezioni sagittali mediane. Nello specifico mostreremo le matrici di confusione dei due test principali: osso mascellare e plesso corioideo. Per l'individuazione della regione contenente la translucenza nucale mostreremo la frequenza della corretta identificazione della regione. Infine, per la misurazione della translucenza nucale evidenzieremo quanto la misurazione effettuata dalla tecnica proposta si discosta, in termini di pixel, dalla misurazione effettuata manualmente dallo specialista accreditato dalla Fetal Medicine Foundation.

4.6.1 Sezione sagittale mediana

Dai filmati in nostro possesso sono estratti ulteriori 2908 fotogrammi casuali, sui quali è stata testata l'intera metodologia. Le matrici di efficienza, qui di seguito esprimono la percentuale dei fotogrammi correttamente classificati dalla prima rete neurale, la quale esegue una classificazione in base alla presenza del ramo verticale dell'osso mascellare. Delle 2908 immagini date in input, 1078 superano il test relativo all'osso mascellare, di cui 847 sono state classificate correttamente non individuando il ramo verticale, nelle restanti 231 il sistema commette un errore di classificazione, pur essendo il ramo verticale presente, il sistema non è in grado di individuarlo. Solo il 5%, ovvero in 45 frames, delle immagini in cui il ramo verticale non è presente vengono scartate dalla metodologia erroneamente (Tabella 4.3).

Tabella 4.3: Matrice delle performance della rete neurale I.

		Medico	
		si	no
Metodologia	si	94.96%	11.46%
	no	5.04%	88.54%

Test ramo verticale

Per completezza, anche se non costituisce un fattore discriminante tra i fotogrammi sottoposti al secondo test, relativo al plesso corioideo, riportiamo i risultati della classificazione per la seconda rete neurale (Tabella 4.4).

Tabella 4.4: Matrice delle performance della rete neurale II.
Medico

		Medico			
		mandibola	mento	naso	altro
Metodologia	mandibola	98.56%	0.20%	0.10%	0.63%
	mento	0.85%	99.80%	0.10%	3.54%
	naso	0.00%	0.00%	99.70%	20.60%
	altro	0.59%	0.00%	0.10%	75.23%

Test componenti facciali

La matrice di confusione in Tabella 4.5 rappresenta l'efficienza della metodologia nel caso del test del plesso corioideo applicato sui 1078 fotogrammi che hanno superato con successo il test dell'osso mascellare.

Tabella 4.5: Matrice delle performance del test sul plesso corioideo.
Medico

		Medico	
		si	no
Metodologia	si	86.48%	16.73%
	no	13.52%	83.27%

Plesso corioideo test

Considerato che, dopo il primo test, parte delle 231 immagini classificate erroneamente come possibili sezioni sagittali mediane saranno comunque respinte dal secondo test relativo al plesso corioideo, per dare una rappresentazione della robustezza dell'intera metodologia, riportiamo in Tabella 4.6 la matrice di confusione relativamente a tutte le 2098 immagini.

Tabella 4.6: Matrice delle performance della metodologia per l'individuazione della sezione sagittale mediana.

		Medico	
		accettata	scartata
Metodologia	accettata	87.26%	5.02%
	scartata	12.74%	94.98%

Test sagittale-mediana

4.6.2 Individuazione della translucenza nucale

La metodologia proposta nel paragrafo “Individuazione della regione nucale” basata sulle relazioni spaziali che legano l'osso mascellare, l'osso frontale e la regione nucale, permette la corretta identificazione nel 99.95% dei casi. Si osserva che saltuariamente la regione nucale viene erroneamente individuata poichè regioni prossime ad essa manifestano caratteristiche morfologiche e di anecogenicità simili, come nel caso della *cisterna magna* nella zona superiore o del *sacco amniotico* nella zona inferiore della translucenza nucale, Figura 4.9.

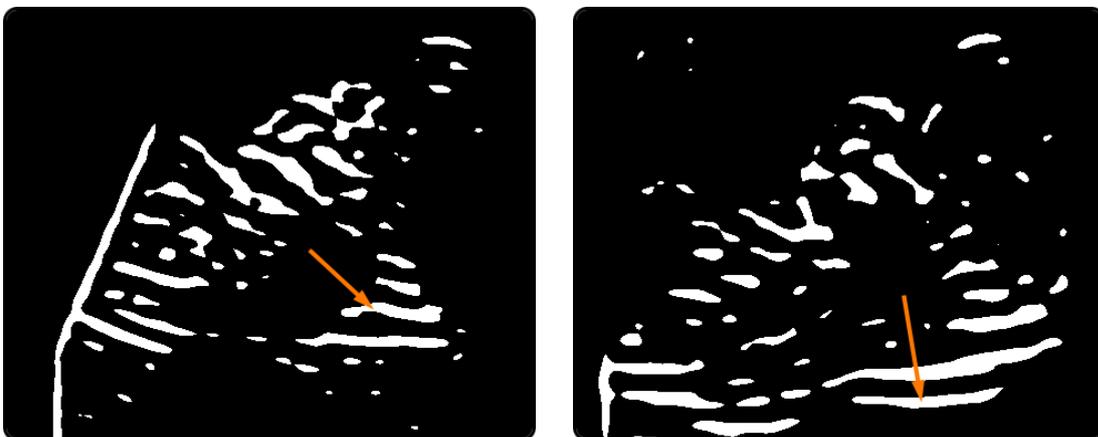


Figura 4.9: La regione anecogena sita superiormente alla translucenza nucale, indicata dalla freccia (sinistra). La regione anecogena sita inferiormente alla translucenza nucale, indicata dalla freccia (destra).

4.6.3 Misurazione della translucenza nucale

In questa sezione sono riportati i confronti effettuati tra le misurazioni eseguite manualmente dall'esperto e le misurazioni effettuate in modo automatico dalla metodologia. Individuata la translucenza nucale, si esprime come sua misura il diametro massimo che congiunge le due componenti ecogene che la delimitano. Non tutti i fotogrammi che rappresentano una sezione sagittale mediana sono candidati ideali per la misurazione della translucenza nucale, in quanto questa può apparire non sufficientemente definita, rendendo difficile il posizionamento del caliper da parte dell'esperto. Il confronto è riportato in termini di pixel, in quanto l'ecografo con il quale sono stati acquisiti i filmati non permette di risalire alle informazioni legate alla risoluzione. Attraverso una interfaccia grafica realizzata appositamente, l'esperto ha posizionato nelle immagini i caliper che delimitano la translucenza e la misurazione è stata registrata considerando il numero di pixel (in direzione verticale) compresi tra i due caliper. Diverse modalità di confronto tra le misurazioni manuale e automatiche sono riportate. Indichiamo con **A** l'insieme di fotogrammi riconosciuti in modo automatico dalla metodologia come sezioni sagittali mediane; dall'insieme **A** l'esperto ha selezionato i fotogrammi più idonei per una migliore misurazione della translucenza nucale, indichiamo con **B** tale insieme.

Poichè a fotogrammi appartenenti allo stesso filmato dovrebbero corrispondere misurazioni con valori molto simili e ipotizzando una minima variabilità dipendente dall'operatore (assumendo di minimizzare la variabilità operatore dipendente), è stato considerato come diametro rappresentativo per ciascuna delle sequenze video, il valore medio delle misurazioni eseguite sui fotogrammi manualmente e automaticamente, e di queste si è calcolata la differenza; il valore medio dell'errore commesso, inteso come differenza tra le misurazioni, è pari a 1.59 pixel.

Per migliorare le prestazioni della metodologia, con la consapevolezza della presenza di outlier nelle misurazioni automatiche, abbiamo deciso di determinare il sottoinsieme **C** di **A**, composto dai fotogrammi in cui diametro è compreso nell'intervallo $[\mu_s - \sigma_s, \mu_s + \sigma_s]$ dove con μ_s e σ_s indichiamo lo spessore medio e la deviazione standard dello spessore della translucenza nucale relativamente alla sequenza video s a cui i fotogrammi appartengono.

Di seguito, i grafici mostrano l'errore commesso dalla metodologia considerando i singoli fotogrammi appartenenti all'insieme **B**. Nel primo l'errore è espresso come la differenza in

valore assoluto tra le misurazioni manuali e automatiche Figura 4.10, nel secondo per ciascun fotogramma (asse delle ascisse) sono riportate le misure manuali e automatiche dei diametri in termini di pixel (asse delle ordinate).

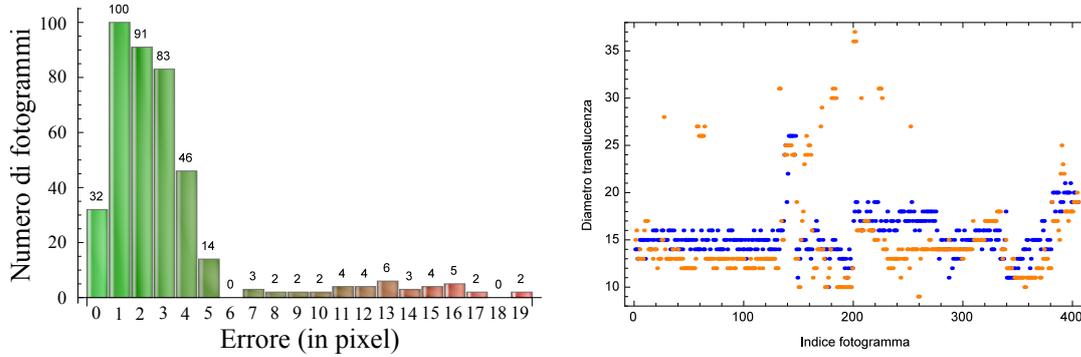


Figura 4.10: Istogramma dell’errore in valore assoluto tra le misurazioni manuale e automatiche per i fotogrammi dell’insieme **B** (sinistra). Per ciascun fotogramma, in ascissa, in rosso il valore delle misurazioni manuali e in blu il valore delle misurazioni automatiche (destra).

L’errore medio calcolato su tutti i fotogrammi è 3.17 con una deviazione standard pari a $\sigma = 3.51$.

Si osserva un miglioramento quando l’insieme **C** dei fotogrammi è considerato, ovvero quando gli outlier sono esclusi dal confronto. Riportiamo i grafici analoghi a Figura 4.10, considerando l’insieme dei fotogrammi appartenenti all’insieme $\mathbf{B} \cap \mathbf{C}$, motivo per cui i grafici sono riportati per un numero di fotogrammi inferiore (Figura 4.11).

L’errore medio calcolato su tutti i fotogrammi è 2.28 con una deviazione standard pari a $\sigma = 1.73$. Per dare una valutazione qualitativa mostriamo alcune immagini (Figura 4.12) nelle quali l’area della traslucenza nucale identificata dalla metodologia è messa in evidenza.

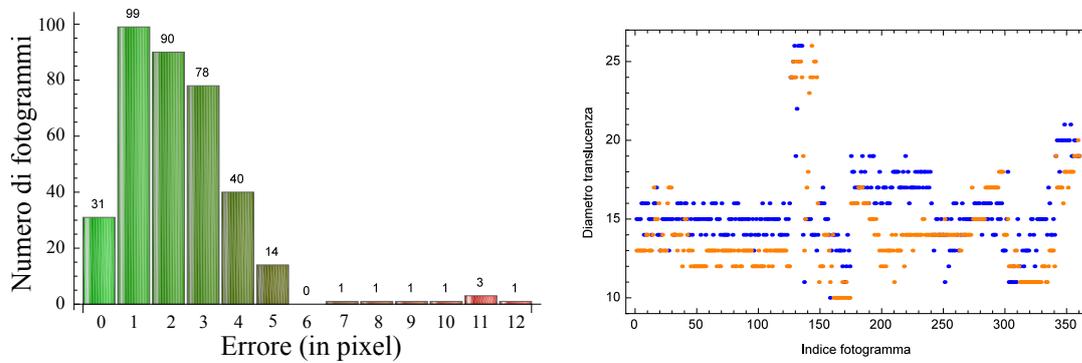


Figura 4.11: Istogramma dell'errore in valore assoluto tra le misurazioni manuale e automatiche per i fotogrammi dell'insieme $\mathbf{B} \cap \mathbf{C}$ (sinistra). Per ciascun fotogramma, in ascissa, in rosso il valore delle misurazioni manuali e in blu il valore delle misurazioni automatiche (destra).

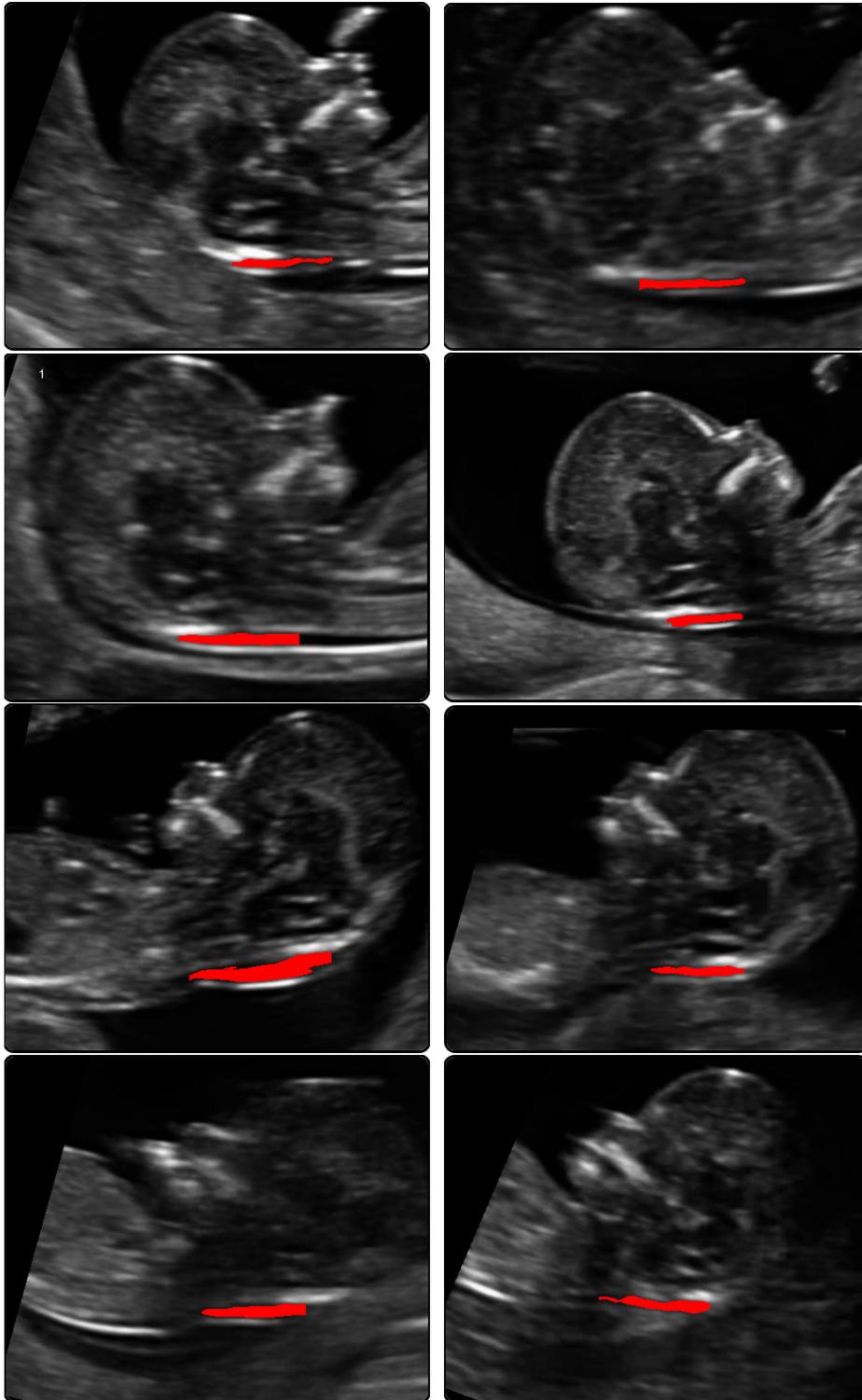


Figura 4.12: In rosso l'area identificata come appartenente alla translucenza nucale.

Bibliografia

- [1] Nicolaides K. H., Azar G., Byrne D., Mansur C., and Marks K. Fetal nuchal translucency: ultrasound screening for chromosomal defects in first trimester of pregnancy. *BMJ*, 304(6831):867–869, 1992.
- [2] Nicolaides K. H., Brizot M. L., and Snijders R. J. M. Fetal nuchal translucency-ultrasound screening for fetal trisomy in the first trimester of pregnancy. *British Journal of Obstetrics and Gynaecology*, 101(9):782–786, 1994.
- [3] Snijders RJM., Noble P., Sebire N., Souka A., and Nicolaides KH. Uk multicentre project on assessment of risk of trisomy 21 by maternal age and fetal nuchal translucency thickness at 10–14 weeks of gestation. *Lancet*, 6:343–351.
- [4] FMF. Fetal Medicine Foundation.
- [5] Wald N. J., George L., Smith D., Densem J. W., and Pettersonm K. On behalf of the International Prenatal Screening Research, G. Serum Screening For Down s Syndrome Between 8 And 14Weeks Of Pregnancy. *BJOG: An International Journal of Obstetrics Gynaecology*, 103(5):407–412, 1996.
- [6] Orlandi F., Rossi C., Orlandi E., Jakil MC., Hallahan TW., Macri VJ., and Krantz DA. First-trimester screening for trisomy-21 using a simplified method to assess the presence or absence of the fetal nasal bone. *Am J Obstet Gynecol*, 192(4):1107–1111, 2005 Apr.
- [7] Krantz DA., Hallahan TW., Orlandi F., Buchanan P., Larsen JW., and Macri JN. First-trimester Down syndrome screening using dried blood biochemistry and nuchal translucency. *Obstet Gynecol*, 96(2):207–213, 2000 Aug.

- [8] Tabor A., Philip J., and Madsen M. Randomised controlled trial of genetic amniocentesis in 4606 low-risk women. *Lancet*, 1:1287–1293, 1986.
- [9] Alfrevic Z., Sundberg K., and Brigham S. Amniocentesis and chorionic villus sampling for prenatal diagnosis. *The Cochrane Database of Systematic Reviews*, Issue 3, 2003.
- [10] Chan Y. T. *Wavelet Basics*. 1995.
- [11] Cohen A., Daubechies I., and Feauveau J.C. Biorthogonal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 45:485–560, 1992.
- [12] Mallat S. *A wavelet tour of signal processing*. 2008.
- [13] Glassner A. S. *Principles of digital image synthesis*. Morgan Kaufmann Publishers, Inc., 1995.
- [14] Gabor D. *Theory of communication*. Institution of Electrical Engineering, 1946.
- [15] Grossmann A. and Morlet J. Decomposition of Hardy functions into square integrable wavelets of constant shape. *Journal of Mathematical Analysis*, 15:723–736.
- [16] Mallat S. "a Theory for Multiresolution Signal Decomposition: The Wavelet Representation". *IEEE Transactions on pattern analysis and machine intelligence*, II(7):674–693, 1989 July.
- [17] Meyer Y. *Wavelets and operators*, Cambridge Studies in Advanced Mathematics. *Cambridge University Press*, 37, 1992.
- [18] Holschneider M., Kronland-Martinet R., Morlet J., and Tchamitchian Ph. "a Real-Time Algorithm for Signal Analysis with the Help of the Wavelet Transform". *Wavelets - Inverse problems and theoretical imaging*. *Springer-Verlag*, pages 286–297, 1990.
- [19] Holschneider M., Kronland-Martinet R., Morlet J., and Tchamitchian Ph. The à trous algorithm. Technical Report CPT-88/P.2215, 1988.
- [20] McCulloch W. S. and Pitts W. A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Analysis*, pages 115–133.

- [21] Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6):386–408.
- [22] Bishop C. M. *Pattern Neural Networks for Pattern Recognition*. Oxford University Press, 1995.
- [23] Minsky M. and Papert S. *Perceptrons: an introduction to computational geometry*. The MIT Press, 1969.
- [24] Bishop C. M. *Pattern Recognition and Machine Learning*. Springer, 2006.
- [25] Vapnik V.N. *The Nature of Statistical Learning Theory*. SpringerVerlag, 1995.
- [26] Vapnik V.N. *Statistical Learning Theory*. Wiley, 1998.
- [27] Loy G. and Zelinsky A. Fast Radial Symmetry for Detecting Points of Interest. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, 25(8), 2003 Aug.
- [28] Reisfeld D., Wolfson H., and Yeshurun Y. Context free attentional operators: the generalized symmetry transform. *International Journal of Computer Vision, Special Issue on Qualitative Vision*, 14:119–130, 1995.
- [29] Intrator N., Reisfeld D., and Yeshurun Y. Extraction of facial features for recognition using neural networks. *Proceeding of the International Workshop on Automatic Face and Gesture Recognition*, 1995.
- [30] Reisfeld D. and Yeshurun Y. Preprocessing of face images: Detection of features and pose normalisation. *Computer Vision and Image Understanding*, 71(3):413–430, 1998 Sept.
- [31] Sela G. and Levine M.D. Real-time attention for robotic vision. *Real-Time Imaging*, 3:173–194, 1997.
- [32] Kimme C., Ballard D., and Sklansky J. Finding circles by an array of accumulators. *Communications of the Association for Computing Machinery*, 18(2):120–122, 1975 Feb.

- [33] Minor L. G. and Sklansky J. Detection and segmentation of blobs in infrared images. *IEEE Transactions on Systems Man and Cybernetics, SMC*, 11(3):194–201, 1981 March.
- [34] Gonzalez R. and Woods R. *Digital Image Processing (3rd Edition)*. Prentice-Hall, Inc., 2008.
- [35] Cicero S., Dezerega V., Andrade E., Scheier M., and Nicolaides K. H. Learning curve for sonographic examination of the fetal nasal bone at 11-14 weeks. *Ultrasound in Obstetrics and Gynecology*, 23(2):135–137, 2003.
- [36] Bernardino F., Cardoso R., Montenegro N., Bernardes J., and De Sa J. M. Semiautomated ultrasonographic measurement of fetal nuchal translucency using a computer software tool. *Ultrasound in Medicine and Biology*, 24(1):51–54, 1998.
- [37] Lee Y.-B., Kim M.J., and Kim M.H. Robust border enhancement and detection for measurement of fetal nuchal translucency in ultrasound images. *Medical and Biological Engineering and Computing*, 45(11):1143–1152, 2007.
- [38] Perona P. and Malik J. Scale-space and edge detection using anisotropic diffusion. *Pattern analysis and machine intelligence. IEEE Transactions On*, 12(7):629–639, 1990.
- [39] Catanzariti E., Fusco G., Isgrò F., Masecchia S., Prevete R., and Santoro M. A semi-automated method for the measurement of the fetal nuchal translucency in ultrasound images. *Image Analysis and Processing-ICIAP*, pages 613–622, 2009.
- [40] Nirmala S. and Palanisamy V. Measurement of nuchal translucency thickness for detection of chromosomal abnormalities using first trimester ultrasound fetal images. *Int. J. Comput. Sci. Inf. Secur.*, 6(3):101–106, 2009.
- [41] Deng Y., Wang Y., and Chen P. Estimating fetal nuchal translucency parameters from its ultrasound image. *Proceeding of the International Conference on Bioinformatics and Biomedical Engineering (ICBBE)*, pages 2643–2646, 2008.
- [42] Deng Y., Wang Y., and Chen P. Automated detection of fetal nuchal translucency based on hierarchical structural model. *Proceeding of the IEEE 23rd International Symposium on ComputerBased Medical Systems (CBMS)*, pages 78–84, 2010.

- [43] Deng Y., Wang Y., Chen P., and Yu J. A hierarchical model for automatic nuchal translucency detection from ultrasound images. *Comput Biol Med*, 42(6):706–713, 2012.
- [44] Moratalla J., Pintero K., Minekawa R., Lachmann R., Wright D., and Nicolaides K. H. Semi-automated system for measurement of nuchal translucency thickness. *Ultrasound in Obstetrics and Gynecology*, 36(4):412–416, 2010.
- [45] Supriyanto Eko, Wee Lai Khin, and Min Too Yuen. "ultrasonic Marker Pattern Recognition and Measurement Using Artificial Neural Network". *WSEAS International Conference on Signal Processing*, pages 35–40, 2010.
- [46] Park J.H., Sofka M., Lee S.M., Kim D.Y., and Zhou S.K. Automatic nuchal translucency measurement from ultrasonography. *Proceeding of International Conference on Medical Image Computing and Computer Assisted Intervention (MICC)*, 2013.
- [47] Sonia R. and Shanthi V. Image Classification for Ultrasound Fetal Images with Increased Nuchal Translucency during First Trimester Using SVM Classifier. *Research Journal of Applied Sciences, Engineering and Technology*, 9(2):113–121, 2015.
- [48] Anzalone A., Fusco G., Isgrò F., Orlandi E., Prevete R., Sciortino G., Tegolo D., and Valenti C. A system for the automatic measurement of the nuchal translucency thickness from ultrasound video stream of the foetus. *IEEE 26th International Symposium on Computer-Based Medical Systems (CBMS)*, pages 239–244, 2013.
- [49] Perona P. and Malik J. Scale-space and edge detection using anisotropic diffusion. *PAMI*, 12:629–639, 1990.
- [50] Rankin R. N., Fenster A., Downey D. B., Munk P. L., Levin M. F., and Vellet A. D. Three dimensional sonographic reconstruction: techniques and diagnostic applications. *American Journal of Roentgenology*, 161(4):695–702, 1993.
- [51] Laudy J. A. M., Janssen M. M. M., Struyk P. C., Stijnen T., Wallenburg H. C. S., and Wladimiroff J. W. Fetal liver volume measurement by three-dimensional ultrasonography: a preliminary study. *Ultrasound in Obstetrics and Gynecology*, 12(2):93–96, 1998.

- [52] Pohls U. G. and Rempfen A. Fetal lung volumetry by three-dimensional ultrasound. *Ultrasound in Obstetrics and Gynecology*, 11(1):6–12, 1998.
- [53] Blaas H. G., Eik-Nes S. H., Berg S., and Torp H. In vivo three-dimensional ultrasound reconstructions of embryos and early fetuses. *Lancet*, 352(9135):1182–1186, 1998.
- [54] Baba K., Okai T., Kozuma S., and Taketani Y. Fetal abnormalities: evaluation with realtime-processible three-dimensional US preliminary report. *Radiology*, 211(2):441–446, 1999.
- [55] Lee A., Deutinger J., and Bernaschek G. Three dimensional ultrasound abnormalities of the fetal face in surface and volume rendering mode. *British Journal of Obstetrics and Gynaecology*, 102(4):302–306, 1995.
- [56] Paul C., Krampfl E., Skentou C., Jurkovic D., and Nicolaides K. H. Measurement of fetal nuchal translucency thickness by three dimensional ultrasound. *Ultrasound in Obstetrics and Gynecology*, 18(5):481–484.
- [57] Lai K.W., Y.C. Hum, and S. Eko. Computerized nuchal translucency three dimensional reconstruction, visualization and measurement for trisomy 21 prenatal early assessment. *Int. J. Phys. Sci.*, 6(19):4640–4648, 2011.
- [58] Cho H.Y., Kwon J.Y., Kim Y.H., Lee K.H., Kim J., Kim S.Y., and Park Y.W. Comparison of nuchal translucency measurements obtained using Volume NT(TM) and two- and three-dimensional ultrasound. *Ultrasound Obst. Gyn.*, 39(2):175–180, 2012.
- [59] Ping C., Nie S., Deng Y., He P., Wang Y., and Yu J. A hierarchical model for automated standard sagittal-view detection from 3d ultrasound data in 11-14. *Biomedical Signal Processing and Control*, 19:96–101, 2015 May.
- [60] Cho HY., Kim YH., Park YW., Kim SY., Lee KH., Yoo JS., and Kwon JY. Image Settings Affecting Nuchal Translucency Measurement Using Volume NT Software. *Yonsei Med J.*, 25(5):1345–1351, 2015 Sep.
- [61] International Telecommunication Union. Recommendation ITU-T H.264 / International Standard ISO/IEC 14496-10. Technical report, 2013.

- [62] Sciortino G., Orlandi E., Valenti C., and Tegolo D. Wavelet analysis and neural network classifiers to detect mid-sagittal sections for nuchal translucency measurement. *Original Research Paper Submitted to Image Analysis & Stereology*. First submission: June 10, 2015. First revision: September 16, 2015.